Computational investigations of ionotropic glutamate receptor ligand binding and conformational change.

by

Alvin Yu

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

August, 2017

Abstract

Ionotropic glutamate receptors (iGluRs) are key mediators of neuronal communication and synaptic plasticity. They are implicated in cellular mechanisms for learning and memory and can be found in the nervous systems of a diverse array of biological organisms from the complex brains in humans to the much simpler nerve nets in ctenophores. At excitatory synapses, these ligand gated ion channels sense the release of glutamate by presynaptic neurons and bind glutamate, triggering conformational change in the ligand binding domains (LBD) that opens the channel pore to allow current flow into the post synaptic neuron. Here we examine the molecular mechanisms of two processes involved in glutamate receptor activation - ligand binding and LBD domain closure. In the AMPA receptor, a glutamate receptor subtype, we present the first simulations of ligand binding in an iGluR and identified metastable binding sites, which facilitate the diffusion of ligand into the binding pocket. We also applied the string method, a chain-of-states pathway sampling method, to investigate AMPA receptor glutamate binding and found that surprisingly, the ligand may bind via at least two distinct pathways. In ctenophore glutamate receptors, early ancestors of NMDA receptors in humans, we calculated the free energy landscape of conformational change to explain puzzling changes in kinetics and thermodynamics of a particularly high affinity glycine-binding domain. These findings both further our understanding of glutamate receptor activation and have broader implications as to the physical and chemical basis for neural activity.

Thesis Advisor: Prof. Albert Lau

Second Reader: Prof. Jeffrey Gray

Thesis Commitee: Prof. Mario L. Amzel, Prof. James M. Berger, Prof Elijah Roberts, Prof. Dominique Frueh, Prof. Jennifer Kavran

Acknowledgments

There are certainly too many people to thank and acknowledge without whom the work here would surely not have been possible. I'd like to begin by acknowledging my family and friends.

To my mother, Annie Yu, whose unending support and life sacrifices have imprinted on me a deep desire to be the best possible person that I can be. To my father, Howard Yu, whose financial support and kind advice has allowed me to chase my dreams, even though it may not be the wisest of financial decisions. To my sisters, Alice and Stephanie Yu, I will probabaly always be the annoying little brother, but their examples and achievements have always served as a guide for me.

To Catherine Ho, who has always provided emotional support and humor when I needed it the most. To John Belcher, whose humor I have appreciated for the past 6 years and still teaches me things about popular culture that I did not know. To Tyler Wied, whose hard work has always caused him to arrive in lab earlier than me. Thanks for opening the door, so I don't get lost in the hallway.

To my academic advisor, Albert Lau, who has provided invaluable insight into the scientific challanges I faced, taught me what science is, and provided me with mentorship and support whenever I needed it. Truly, he has taught me to become a better scientist.

To my thesis committee: Mario Amzel for always giving me thought-provoking discussions and teaching me about the NIS protein. To James Berger, Jeffrey Gray, Elijah Roberts, Dominique Frueh, and Jennifer Kavran for their insightful comments.

To all these people and more, I would like to express my deepest gratitude, and note that my career at Johns Hopkins would not be possible without.

Contents

List of Figures

List of Tables

Chapter 1 - Introduction

1.1 iGluR structure

Ionotropic glutamate receptors are ion channels comprised of four subunits that surround a central channel pore. Each subunit corresponds to an iGluR subtype, which can be grouped into distinct classes based on their pharmacological properties: the AMPA receptors with subtypes GluA1-GluA4, the Kainate receptors with subtypes GluK1-GluK5, the NMDA receptors with subtypes GluN1, GluN2A-GluN2D, GluN3A, and GluN3B, and the δ receptors with subtypes GluD1 and GluD2 [1]. Functional receptors are formed only with a subunit composition within the same glutamate receptor class. Different classes are thought to play slightly different roles in the complex mechanisms responsible for synaptic transmission. All the iGluRs have a modular architecture in which the subunits are assembled in a "dimer-of-dimers" fashion [2], and each subunit can be divided into three separate domains, an amino-terminal domain (ATD), ligand-binding domain (LBD), transmembrame domain (TMD) and C-terminal domain (CTD) **Fig. 1-1**. For the remainder of this introduction, we will focus on only two of the iGluR classes: AMPA and NMDA receptors.

The AMPA-type glutamate receptor can form both homo- and heteromers with its varying subunits (GluA1-GluA4). The first crystal structure of an AMPA-type iGluR homomer was solved by Sobolevsky et al. in 2009 [3], and more recently, the heteromeric structure of an AMPA-type iGluR was solved by Herguedas et al. in 2016 [4]. Both structures revealed that, broadly speaking, AMPA receptors are two-fold symmetric and contain only two pairs of conformationally distinct subunits, subunits AC and subunits BD, with subunit A being equivalent to subunit C and subunit B being equivalent to subunit D. Conformational change in one region of the receptor can be propagated throughout the entire receptor; in the LBD layer, subunit A,B and subunit C,D form dimers, whereas in the ATD layer, subunit A,D and subunit B,C form dimers. The binding of glutamate triggers conformational change in the LBD layer, which is propagated to the TMD layer through the LBD - TMD linkers, and opens the ion channel pore. The ATD layer is thought to be

involved in the assembly of the tetrameric receptor and receptor trafficking [5–7], but is nonessential to receptor function - AMPA receptor mutants that lack the ATD still form functional channels [8, 9]. AMPA receptors are primarily responsible for fast excitatory transmission in the central nervous system.

The NMDA-type glutamate receptors, on the other hand, are obligate heteromers and require two GluN1 subunits to form functional channels. NMDA tetramers can be composed of either two GluN1 subunits with two GluN2 subunits or a combination of GluN2 and GluN3 subunits. Unlike AMPA receptors, which require only glutamate to activate, NMDA receptors require the binding of both glutamate and a co-agonist, glycine, to activate. The LBDs of GluN2 subunits serve as the binding site for glutamate, whereas the GluN1 and GluN3 LBDs subunits contain binding sites for glycine. Glycine-activated NMDA receptor channels can be formed by tetramers of only GluN1 of GluN3 subunits [10]; these channels have not been found to be present in neurons [11]. One notable distinction of NMDA receptors is that under physiological conditions, the channel pore is blocked by a Mg^{2+} ion, which is only released once the membrane is depolarized. Thus, the majority of NMDA receptors exhibit a voltage dependence and require both ligand binding and membrane depolarization to activate. Only AMPA receptors which lack the GluA2 subtype have a functional voltage dependence - intracellular polyamines can serve as channel blockers in the AMPA receptor TMD. In AMPA receptors, this polyamine block occurs after the membrane is depolarized, as opposed to the Mg^{2+} block in NMDA receptors, which is released once the membrane is depolarized. NMDA receptors, like AMPA receptors, are two-fold symmetric and can function without a proper ATD [12] and adopts a "ballon-shape" conformation rather than the characteristic Y-shaped conformation of GluA2 homomers.

1.2 Molecular Simulations of iGluRs

Computer simulations of the fundamental interactions required to model the behavior of a system has a rich history. The first numerical simulations were performed by Enrico Fermi, John Pasta, and Stanislaw Ulam in 1955 on a one-dimensional vibrating

string of particles [13]. Aneesur Rahman published the first simulations of liquid argon using a Lennard-Jones potential in 1964 [14]. Finally, the first molecular dynamics simulations of a biological macromolecule was published by J. Andrew McCammon, Bruce R. Gelin, and Martin Karplus in 1977 on the bovine pancreatic trypsin inhibitor [15]. Today, molecular dynamics is a commonly used computational tool to probe the behavior of chemical and biological systems.

Typically in a molecular dynamics simulation, Newton's equations of motion: $F = -\nabla U$ are solved for every particle in the system under consideration. The algorithm responsible for the numerical integration of the equations of motion is termed the integrator. Special consideration should be paid to the interatomic potentials or "force-fields" used as they often determine the range and accuracy of the physical phenomena to be studied. Furthermore, the size of the system determines the computational cost of a particular simulation. Currently the most computationally intensive task is the evaluation of pair-wise electrostatic interactions, which scales as $O(n^2)$; although particle mesh Ewald algorithms reduce this to $O(n \log n)$.

Molecular simulations of iGluRs have proven useful in determining the properties of these receptors. X-ray crystallographic analyses have shown that iGluR LBDs adopt two distinct structural conformations, open in the apo form, and closed once agonist is bound [16, 17]. MD simulations of the structural fluctuations of AMPA-type iGluR LBDs spanning $\sim$2 ns demonstrated hints of domain closure motions in the apo form in 2002 [18]. Electrostatic calculations of the apo LBD found that molecular surface of the binding site contains a mostly positive electrostatic potential which may attract negatively charged ligands [19]. The free energy of ligand binding was calculated for several different agonists [20] as well as the free energy landscape of domain closure motions [21, 22]. Furthermore, the LBD-TMD linkers couples the useful work performed during domain closure to opening of the channel pore [23].

1.3 Central Objectives

The central objective of this dissertation is to understand the molecular processes governing iGluR behavior. Like other ion channels, iGluRs undergo both activation and

densensitization. During activation, chemical stimuli in the form of glutamate binding leads to the rapid conduction of ions. LBD domain closure and other conformational rearrangements are proposed to play a role [24]; although the precise molecular mechanisms involved are still unclear. Desensitization occurs when further application of stimuli results in submaximal or no current response from the receptor. In AMPA receptors, desensitization involves conformational rearrangements which disrupt the LBD dimer interface [25]. The molecular details of AMPA receptor densensitization, however, are not well understood.

AMPA receptors activate within a few hundred μs and densensitize in ms whereas NMDA receptors have slower kinetics [26]. The fast kinetics of AMPA receptors suggest that AMPA receptor processes are particularly amenable to MD simulations. NMDA receptor processes on the other hand, may require greater degrees of extrapolation or more focused simulations which examine the mechanisms piecemeal. The content of this dissertation is focused on two molecular processes which occur during iGluR activation - ligand binding, and LBD conformational change.

1.3.1 Aim 1: iGluR ligand binding

Crystal structures of the agonist bound AMPA-type LBD feature several protein-ligand interactions in the binding pocket [17]. The glutamate ligand is orientated with the α-carboxyl group pointing towards Lobe 1 of the LBD, anchored by a positively charged arginine residue (R485). Interactions between R485 and glutamate are essential for binding - R485 is conserved across all glutamate receptor classes, and mutations of the residue to lysine abolish binding in GluA1 and GluA4 LBDs [27,28]. The ligand's amide group contacts a negatively, charged glutamate (E705) side chain at the center of the binding pocket, and forms additional hydrogen-bonding interactions with Thr 480 and Pro 478. The ligand's γ-carboxyl group is further secured by interactions with the amide backbones of Ser 654 and Thr 655 on Lobe 2 of the LBD. Prior to our studies, time-resolved fourier-transform infrared spectroscopy (FTIR) experiments suggested that glutamate binds through a two step mechanism involving first, establishment of α-carboxyl interactions with Lobe 1 and secondly, domain clo-

sure and formation of γ-carboxyl interactions with residues on Lobe 2 of the LBD [29]. Despite much success in determining the free energies of ligand binding [20] and capturing structural conformations of ligand bound states, little is known of the molecular mechanisms by which the ligand finds its way into the binding pocket. Modelling this process is one of the primary goals of this dissertation.

1.3.2 Aim 2: iGluR LBD conformational change

LBDs close in a clamshell-like motion in response to agonist binding. Crystallographic structures of the glutamate bound and apo AMPA type LBDs have detected an $\sim 20^{o}$ difference between Lobe 1 and Lobe 2 of the LBD, as measured along an axis of rotation parallel to the to the axis of helix I and perpendicular to the two interdomain β-strands. Free energy landscapes of AMPA-receptor domain closure motions have previously been calculated and found to contribute substantially to the free energy of ligand binding [20], demonstrating that much of the useful work performed by ligand binding is directly translated into domain closure motions. The energetics of domain closure motions have also been computed for other iGluR classes, including the NMDA receptor [22]. However, several open questions remain regarding LBD conformational change. How is ligand binding mechanistically coupled to domain closure? To what degree can domain closure motions be tuned by specific interactions between Lobe 1 and Lobe 2? How does LBD domain closure induce pore opening? We attempt to address the former two questions in this disseration, and leave the latter for future studies.

1.4 Overview of remainder of thesis

The remainder of this thesis is motivated by the two aims above, and will be organized as follows.

Chapter 2 provides a brief overview of molecular dynamics simulations, both theory and implementation for one of the commonly used free energy methods used in this thesis (e.g. umbrella sampling), and instructions on the application of free energy methods to iGluRs.

Chapter 3 explores the energetics of LBD conformational change of a particularly high affinity glycine-binding domain, which is evolutionarily related to glycine binding NMDA subunits (GluN1/GluN3). The theoretical free energy landscapes calculated here provide the basis for understanding how a single interdomain interaction, between an arginine on Lobe 1 and glutamate on Lobe 2, may tune the LBD binding affinity by three orders of magnitude and densensitization kinetics of the receptor by two orders of magnitude.

Chapter 4 presents the first simulations of ligand binding in an AMPA-type iGluR LBD. Metastable interactions with positively charged residues on the surface of the LBD are found to funnel the negatively charged ligand into the binding pocket. 3D free energy landscapes that calculate the energetics of transient binding sites for the ligand are used to delineate the possible pathways with which the ligand may bind. Electrophysiological recordings demonstrate that elimination of these transient binding sites significantly affects the activation kinetics of the receptor.

Chapter 5 examines ligand binding in the AMPA-type iGluR LBD using a "chain-of-states" technique, the string method, in which the ligand binding process is split into individual transition states between the ligand bound and apo conformations. This presents several computational advantages over brute force simulation methods, including greater sampling of the transition event, which is particularly useful if the transition event is rare, as in ligand binding. Multiple pathways for ligand binding were found to be possible. The energetics of different pathways were evaluated.

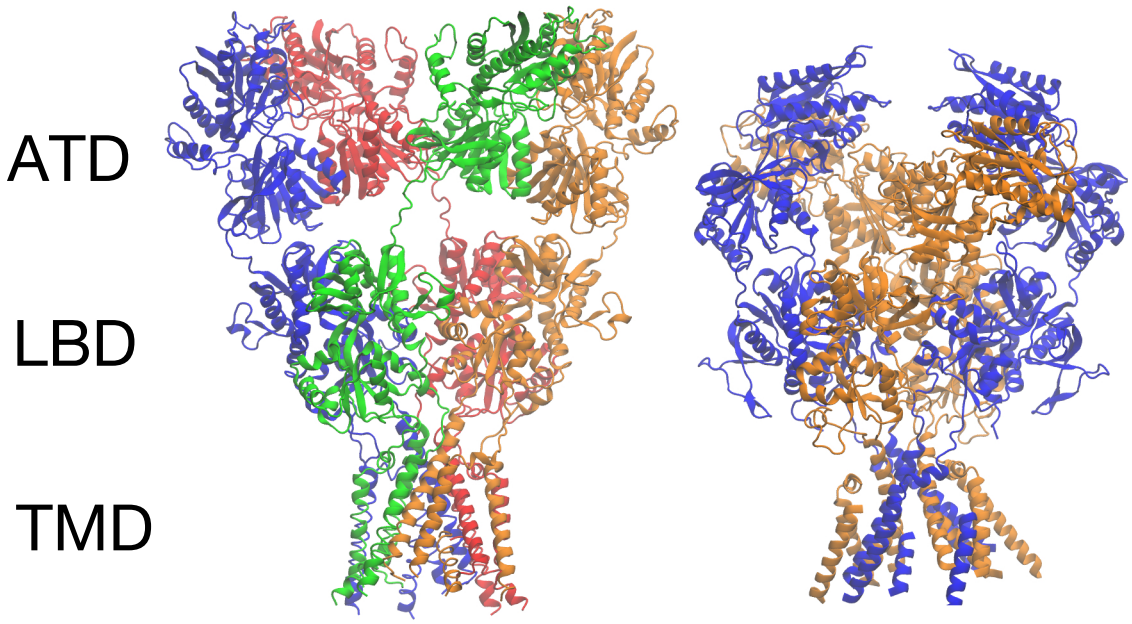

Fig. 1-1: **Full-length AMPA and NMDA receptors.** The crystal structure of the intact AMPA homomer (PDB ID: 3KG2), left panel. Subunits A, B, C, D are shown in blue, red, green, and orange respectively. Subunit A is equivalent to subunit C, whereas subunit B is equivalent to subunit D. The crystal structure of the intact NMDA heterotetramer (PDB ID: 4TLL) composed of two GluN1 and two GluN2B subunits, right panel. GluN1 subunits which bind glycine are shown in blue, whereas GluN2B subunits which bind glutamate are shown in orange. Both AMPA and NMDA receptors can be separated into four different layers: the ATD, LBD, TMD, and CTD.

References

[1] Traynelis, S. F., Wollmuth, L. P., McBain, C. J., Menniti, F. S., Vance, K. M., Ogden, K. K., Hansen, K. B., Yuan, H., Myers, S. J., and Dingledine, R. Glutamate receptor ion channels: structure, regulation, and function. *Pharmacological Reviews*, 62(3):405–496, September 2010.

[2] Mayer, M. L. Structural biology of glutamate receptor ion channel complexes. *Current Opinion in Structural Biology*, 41:119–127, December 2016.

[3] Sobolevsky, A. I., Rosconi, M. P., and Gouaux, E. X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature*, 462(7274):745–756, December 2009.

[4] Herguedas, B., GarcÃŋa-NafrÃŋa, J., Cais, O., FernÃąndez-Leiro, R., Krieger, J., Ho, H., and Greger, I. H. Structure and organization of heteromeric AMPA-type glutamate receptors. *Science (New York, N.Y.)*, 352(6285):aad3873, April 2016.

[5] Ayalon, G. and Stern-Bach, Y. Functional assembly of AMPA and kainate receptors is mediated by several discrete protein-protein interactions. *Neuron*, 31(1):103–113, July 2001.

[6] Mansour, M., Nagarajan, N., Nehring, R. B., Clements, J. D., and Rosenmund, C. Heteromeric AMPA receptors assemble with a preferred subunit stoichiometry and spatial arrangement. *Neuron*, 32(5):841–853, December 2001.

[7] Ayalon, G., Segev, E., Elgavish, S., and Stern-Bach, Y. Two regions in the N-terminal domain of ionotropic glutamate receptor 3 form the subunit oligomerization interfaces that control subtype-specific receptor assembly. *The Journal of Biological Chemistry*, 280(15):15053–15060, April 2005.

[8] Pasternack, A., Coleman, S. K., Jouppila, A., Mottershead, D. G., Lindfors, M., Pasternack, M., and KeinÃď'nen, K. Alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor channels lacking the N-terminal do-

main. *The Journal of Biological Chemistry*, 277(51):49662–49667, December 2002.

[9] Horning, M. S. and Mayer, M. L. Regulation of AMPA receptor gating by ligand binding core dimers. *Neuron*, 41(3):379–388, February 2004.

[10] Chatterton, J. E., Awobuluyi, M., Premkumar, L. S., Takahashi, H., Talantova, M., Shin, Y., Cui, J., Tu, S., Sevarino, K. A., Nakanishi, N., Tong, G., Lipton, S. A., and Zhang, D. Excitatory glycine receptors containing the NR3 family of NMDA receptor subunits. *Nature*, 415(6873):793–798, February 2002.

[11] Matsuda, K., Fletcher, M., Kamiya, Y., and Yuzaki, M. Specific assembly with the NMDA receptor 3b subunit controls surface expression and calcium permeability of NMDA receptors. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 23(31):10064–10073, November 2003.

[12] Yuan, H., Hansen, K. B., Vance, K. M., Ogden, K. K., and Traynelis, S. F. Control of NMDA receptor function by the NR2 subunit amino-terminal domain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29(39):12045–12058, September 2009.

[13] Berman, G. P. and Izrailev, F. M. The Fermi-Pasta-Ulam problem: Fifty years of progress. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 15(1):015104, March 2005.

[14] Rahman, A. Correlations in the Motion of Atoms in Liquid Argon. *Physical Review*, 136(2A):A405–A411, October 1964.

[15] McCammon, J. A., Gelin, B. R., and Karplus, M. Dynamics of folded proteins. *Nature*, 267(5612):585–590, June 1977.

[16] Armstrong, N., Sun, Y., Chen, G. Q., and Gouaux, E. Structure of a glutamate-receptor ligand-binding core in complex with kainate. *Nature*, 395(6705):913–917, October 1998.

[17] Armstrong, N. and Gouaux, E. Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron*, 28(1):165–181, October 2000.

[18] Arinaminpathy, Y., Sansom, M. S. P., and Biggin, P. C. Molecular dynamics simulations of the ligand-binding domain of the ionotropic glutamate receptor GluR2. *Biophysical Journal*, 82(2):676–683, February 2002.

[19] Speranskiy, K. and Kurnikova, M. On the Binding Determinants of the Glutamate Agonist with the Glutamate Receptor Ligand Binding Domain. *Biochemistry*, 44(34):11508–11517, August 2005.

[20] Lau, A. Y. and Roux, B. The hidden energetics of ligand binding and activation in a glutamate receptor. *Nature Structural & Molecular Biology*, 18(3):283–287, March 2011.

[21] Lau, A. Y. and Roux, B. The Free Energy Landscapes Governing Conformational Changes in a Glutamate Receptor Ligand-Binding Domain. *Structure (London, England : 1993)*, 15(10):1203–1214, October 2007.

[22] Yao, Y., Belcher, J., Berger, A. J., Mayer, M. L., and Lau, A. Y. Conformational analysis of NMDA receptor GluN1, GluN2, and GluN3 ligand-binding domains reveals subtype-specific characteristics. *Structure (London, England: 1993)*, 21(10):1788–1799, October 2013.

[23] Kazi, R., Dai, J., Sweeney, C., Zhou, H.-X., and Wollmuth, L. P. Mechanical coupling maintains the fidelity of NMDA receptor-mediated currents. *Nature Neuroscience*, 17(7):914–922, July 2014.

[24] Lau, A. Y., Salazar, H., Blachowicz, L., Ghisi, V., Plested, A. J., and Roux, B. A Conformational Intermediate in Glutamate Receptor Activation. *Neuron*, 79(3), August 2013.

[25] Meyerson, J. R., Kumar, J., Chittori, S., Rao, P., Pierson, J., Bartesaghi, A., Mayer, M. L., and Subramaniam, S. Structural mechanism of glutamate receptor activation and desensitization. *Nature*, 514(7522):328–334, October 2014.

[26] Clements, J. D., Feltz, A., Sahara, Y., and Westbrook, G. L. Activation kinetics of AMPA receptor channels reveal the number of functional agonist binding sites. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 18(1):119–127, January 1998.

[27] Kawamoto, S., Uchino, S., Xin, K. Q., Hattori, S., Hamajima, K., Fukushima, J., Mishina, M., and Okuda, K. Arginine-481 mutation abolishes ligand-binding of the AMPA-selective glutamate receptor channel alpha1-subunit. *Brain Research. Molecular Brain Research*, 47(1-2):339–344, July 1997.

[28] Lampinen, M., PentikÃd'inen, O., Johnson, M. S., and KeinÃd'nen, K. AMPA receptors and bacterial periplasmic amino acid-binding proteins share the ionic mechanism of ligand recognition. *The EMBO journal*, 17(16):4704–4711, August 1998.

[29] Cheng, Q., Du, M., Ramanoudjame, G., and Jayaraman, V. Evolution of glutamate interactions during binding to a glutamate receptor. *Nature Chemical Biology*, 1(6):329–332, November 2005.

Chapter 2 - Computing Conformational Free Energies of iGluR Ligand-Binding Domains

Reproduced in part from:

Ionotropic Glutamate Receptor Technologies, Neuromoethods vol. 106 chap 9

DOI 10.1007/978-1-4939-2812-5_9, Springer Science+Business Media New York

2016

by

Alvin Yu, Tyler Wied, John Belcher, and Albert Y. Lau

Ionotropic glutamate receptors (iGluRs) transduce chemical signals at synapses into electrical impulses. This function relies on concerted conformational changes that are propagated among the linked domains of the tetrameric protein assembly making up each receptor. A key conformational change is the closure of the ligand-binding domain (LBD) upon agonist binding, which eventually gates the transmembrane ion channel domain. The free energy that becomes available for gating transitions is governed by the LBD free energy landscapes for apo and ligand-bound states. These landscapes describe the thermodynamic equilibrium among various LBD conformations. Delineating these landscapes is essential for understanding the molecular driving forces underlying iGluR function. Molecular dynamics free energy simulations offer a means for estimating these quantities, which are difficult to extract from experimental results alone. Here, we describe the process of carrying out a free energy computation using an umbrella sampling strategy for characterizing large-scale conformational changes in iGluR LBDs.

2.1 Background

Structural studies of ionotropic glutamate receptors (iGluRs) have provided valuable insights into the molecular mechanics of activation and desensitization [1]. These studies have helped characterize the conformational changes that take place within an isolated ligand-binding domain (LBD) upon binding of agonists and antagonists in addition to the molecular rearrangements that are associated with desensitization. The numerous crystal structures of isolated iGluR domains and intact iGluR receptors also provide atomistic models that enable molecular simulation studies. Such studies allow one to probe conformational energetics not easily accessed by experimental approaches alone by, for example, considering conformational states that are transient, not heavily populated, or refractory to crystallization. Such considerations are important for understanding iGluRs, which are allosteric receptors that rely on complex molecular dynamics for executing their function.

Straightforward, unbiased molecular dynamics (MD) simulations have shed light on important dynamic events involving the LBD; for example, by suggesting pos-

sible pathways for conformational changes related to receptor activation [2, 3] and rearrangements in the configuration of waters in the binding cleft [4]. While these types of simulations are useful, they can have limitations when applied to complex macromolecular systems that exhibit significant conformational changes. Instead, one can employ a simulation strategy that generates a "free energy landscape", or equivalently, a "potential of mean force" (PMF), along specifically chosen coordinates to obtain a reduced description of conformational changes in the system. The free energy difference between conformational states reflects the probability of finding the system in those states.

This chapter focuses on the computation of free energy landscapes using umbrella sampling MD simulations. The systems of interest here are iGluR LBDs. Due to space constraints, we limit the discussion of numerous general aspects of setting up and carrying out an MD simulation. These topics, however, are covered elsewhere in excellent overviews, e.g., by Cheng and Ivanov [5] and the NAMD tutorials at www.ks.uiuc.edu/Training/Tutorials/namd-index.html. The theory of umbrella sampling is discussed in some detail for the sake of completeness and because of its practical utility, as it provides a guide for writing the computer code necessary for carrying out the calculations. We hope that this chapter will be useful to readers in facilitating computational studies of iGluR conformational energetics.

2.2 Materials

The most widely used molecular dynamics (MD) simulation packages for biomolecules are CHARMM [6], AMBER [7], GROMOS [8], NAMD [9], and GROMACS [10]. CHARMM and AMBER have been in continuous development for the longest, offer advanced sampling and analysis capabilities, and are associated with their own force fields for proteins, nucleic acids, lipids, and carbohydrates [5]. All LBD umbrella sampling simulations described below were performed using CHARMM, although other packages can be used as well. NAMD and GROMACS have been developed to excel in performance (i.e., computational speed), and they support numerous force fields. The package Desmond [11], from D.E. Shaw Research, also offers high performance

and supports multiple force fields. For a comparison of different force fields, see a review by Guvench and MacKerell [12]. VMD [13] is a widely used molecular graphics program for structure and trajectory analysis.

MD packages generally require an operating system based on Unix/Linux. Some aspects of setting up a simulation system may be accomplished on a workstation, but equilibration and production simulations are typically performed on either commodity computing clusters containing many processors or high-end computing platforms available through resources such as the Extreme Science and Engineering Discovery Environment (XSEDE; www.xsede.org).

2.3 Methods

2.3.1 Molecular Dynamics Simulations

At the heart of a molecular mechanics simulation is the potential energy function that relates the atomic coordinates of a system, **R**, to the energy, U [14] (**Fig. 2-1**). The minimal set of energy terms commonly used in biomolecular simulations is the following:

$$U(\mathbf{R}) = \sum_{\text{bonds}} K_b(b-b_0)^2 + \sum_{\text{angles}} K_\theta(\theta-\theta_0)^2 + \sum_{\text{dihedrals}} K_\phi[1+\cos(n\phi-\delta)] + \sum_{\substack{\text{impropers}\\ \text{dihedrals}}} K_\chi(\chi-\chi_0)^2 + \sum_{\text{bonded}} \left\{ \frac{q_i q_j}{\epsilon r_{ij}} + \epsilon_{ij}\left[\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{6}\right]\right\} \quad (2.1)$$

where the equilibrium values for the bond distance, b, the valence angle, θ, and the improper dihedral angle (out-of-plane distortion), χ, are b_0, θ_0, and χ_0, respectively; the multiplicity and phase for the dihedral angle, ϕ, are n and δ; q_i and q_j are the partial atomic charges on atoms i and j; the Ks are the force constants; ϵ is the dielectric constant; r_{ij} is the distance between atoms i and j; and the Lennard-Jones potential well depth and minimum interaction radius are ϵ_{ij} and $R_{\min,ij}$. The "bonded" terms

are the bonds, angles, dihedrals, and improper dihedrals; the "nonbonded" terms are the electrostatic interactions treated using a Coulombic term and the interatomic repulsion and dispersion interactions treated using a Lennard-Jones "6-12" term. These terms are common to most modern biomolecular force fields, including CHARMM, AMBER, GROMOS, and OPLS [15], among others. To increase the accuracy of force fields, additional or alternate energy terms have been incorporated. An example of an additional term is the two-dimensional dihedral energy grid "correction map" (CMAP) applied to protein backbone ϕ and ψ angles in CHARMM [16]. Other examples are described in a review by MacKerell [14]. Once $U(\mathbf{R})$ is defined, the time evolution of a system of atoms may be obtained by integrating Newton's equations of motion,

$$F(\mathbf{R}(t)) = -\nabla U(\mathbf{R}(t)) = M\ddot{\mathbf{R}}(t) \tag{2.2}$$

where t is time, F is the ensemble of forces on the atoms, ∇U is the gradient of the potential, M is the matrix that specifies the mass of every atom, and $\ddot{\mathbf{R}}$ is the second derivative of the atomic coordinates with respect to time. Combining Taylor series expansions for $\mathbf{R}(t+\delta t)$ and $\mathbf{R}(t-\delta t)$ yield the Verlet integration algorithm, which is used to generate the dynamical trajectory of the system. Details concerning the propagation of atomic velocities and positions, as well as many practical considerations regarding MD simulations, are covered in a review by Cheng and Ivanov [5].

2.3.2 Order Parameters

How might one go about computing the change in free energy for a conformational transition in a system? A first step would be to identify a continuous coordinate ξ, also referred to as an order parameter, which is a function of a few or more degrees of freedom in the system, and traverses the system's conformational states of interest. Order parameters along the most important degrees of freedom are often chosen as a reduced descriptor of the system. Any choice of order parameter is possible, such as a distance, an angle, or a root mean square deviation from a reference state. The order parameter may be multidimensional, e.g., two distances or a distance plus an

angle. Choosing an appropriate order parameter for a given system often requires a mix of physical intuition and prior knowledge of the chemical process under consideration. Several novel strategies have been developed to quantitatively determine order parameters [17]. These techniques include combining MD simulations with either principal component analysis (PCA) or a determination of the atomic fluctuations that correspond to the largest amounts of information transfer in a conformational transition. Normal mode analysis (NMA) has also been suggested as a means for identifying the most important collective motions [18–20].

2.3.3 Umbrella Sampling

Once the order parameter, ξ, is chosen, the average distribution function along ξ can be provided by a Boltzmann-weighted average [21, 22]:

$$\langle \rho(\xi) \rangle = \int \frac{\delta[\xi'(\mathbf{R}) - \xi] e^{-U(\mathbf{R})/k_B T} d\mathbf{R}}{\int e^{-U(\mathbf{R})/k_B T} d\mathbf{R}} \tag{2.3}$$

where $U(\mathbf{R})$ is the energy of the system as a function of the atomic coordinates $\mathbf{R}$, $\delta[\cdot]$ is the Dirac delta function, k_B is Boltzmann's constant, and T is temperature. The free energy, or PMF, $W(\xi)$, along ξ can be computed as follows:

$$W(\xi) = W(\xi^*) - k_B T \log \left[\frac{\langle \rho(\xi) \rangle}{\langle \rho(\xi^*) \rangle} \right] \tag{2.4}$$

where ξ^* is an arbitrary reference value. In principle, $W(\xi)$ could be computed directly from a "brute force" molecular dynamics simulation of the system. In practice, however, this approach is often limited because large free energy barriers along ξ can impede the amount of conformational space that is sampled within an allotted simulation time (**Fig. 2-2**). Poor sampling would yield unreliable statistics from which to calculate free energy differences. Consequently, numerous sampling strategies have been developed to more effectively calculate a PMF from molecular dynamics simulation trajectories. One of these strategies is the umbrella sampling method developed by Torrie and Valleau [23]. This method typically involves carrying out multiple independent simulations, where each simulation takes place in the presence of an applied

potential, $w(\xi)$. This artificial bias for each simulation, or window, confines sampling to a region near a chosen value of ξ. The chosen ξ is different for each window. A convenient form for the artificial bias in the ith window is a harmonic function, i.e., $w_i = \frac{1}{2}K(\xi - \xi_i)^2$, where K is the force constant. The set of ξ_i spans the conformational space of interest.

The biased distribution function generated by the ith window is

$$\langle \rho(\xi) \rangle_{(i)} = \langle \rho(\xi) \rangle \frac{e^{-w_i(\xi)/k_B T}}{\langle e^{-w_i(\xi)/k_B T} \rangle} \tag{2.5}$$

where $\langle \rho(\xi) \rangle$ is given by 2.3. The unbiased PMF from the ith window is

$$W_i(\xi) = W(\xi^*) - k_B T \log \left[\frac{\langle \rho(\xi) \rangle_{(i)}}{\langle \rho(\xi^*) \rangle} \right] - w_i(\xi) + F_i \tag{2.6}$$

where F_i is the free energy associated with introducing the window potential, $w_i(\xi)$,

$$e^{-F_i/k_B T} = \langle e^{-w_i(\xi)/k_B T} \rangle \tag{2.7}$$

The process of unbiasing and recombining the results of sampling in all the windows in order to obtain $W(\xi)$ involves calculating F_i. Numerous approaches have been developed to accomplish this task [21–24]. We will discuss only one of these approaches, the weighted histogram analysis method (WHAM) [25], although other valid approaches are available. Essentially, the WHAM estimates the unbiased distribution function by computing a weighted sum of all the sampling data and determining the functional form of the weight factors that minimizes the statistical error [21]. Given N biased window simulations [25, 26],

$$\langle \rho(\xi) \rangle = \sum_{i=1}^{N} [\langle \rho(\xi) \rangle]_{(i),\text{unbiased}} \times \left[\frac{n_i \; e^{-[w_i(\xi) - F_i]/k_B T}}{\sum_{j=1}^{N} n_j \; e^{-[w_j(\xi) - F_j]/k_B T}} \right] \tag{2.8}$$

where $[\langle \rho(\xi) \rangle]_{(i),\text{unbiased}}$ are the individual unbiased distribution functions, and n_i are the number of sampling data points used to generate each biased distribution

function. Using 2.5 and 2.7,

$$[\langle\rho(\xi)\rangle]_{(i),\text{unbiased}} = \langle\rho(\xi)\rangle_{(i)}\ e^{w_i(\xi)/k_BT}\ e^{-F_i/k_BT} \tag{2.9}$$

2.8 can now be rewritten [25, 27]:

$$\langle\rho(\xi)\rangle = \sum_{i=1}^{N} n_i\ \langle\rho(\xi)\rangle_{(i)} \times \left[\sum_{j=1}^{N} n_j\ e^{-[w_j(\xi)-F_j]/k_BT}\right]^{-1} \tag{2.10}$$

Solving 2.10 requires the constants F_j, which can be obtained from a reformulation of 2.7:

$$e^{-F_i/k_BT} = \int \langle\rho(\xi)\rangle e^{-w_i(\xi)/k_BT} d\xi \tag{2.11}$$

2.10 and 2.11, known as the WHAM equations, are interdependent (i.e., 2.10 requires F_j, and 2.11 requires $\langle\rho(\xi)\rangle$) and are therefore solved in a self-consistent manner via an iterative procedure. First, an estimate for the unbiased distribution $\langle\rho(\xi)\rangle$ is obtained using 2.10 and an initial guess for all of the F_j. Next, 2.11 is evaluated using the estimate for $\langle\rho(\xi)\rangle$. The cycle is repeated using the new estimates for F_j until both equations have converged. The equations above are written in terms of a one-dimensional ξ. The approach can be applied to multidimensional scenarios as well. For example, 2.10 and 2.11 can be reformulated as follows for handling a two-dimensional order parameter involving the variables ξ_1 and ξ_2:

$$\langle\rho(\xi_1,\xi_2)\rangle = \sum_{i=1}^{N} n_i\ \langle\rho(\xi_1,\xi_2)\rangle_{(i)} \times \left[\sum_{j=1}^{N} n_j\ e^{-[w_j(\xi_1,\xi_2)-F_j]/k_BT}\right]^{-1} \tag{2.12}$$

and

$$e^{-F_i/k_BT} = \int\int \langle\rho(\xi_1,\xi_2)\rangle e^{-w_i(\xi_1,\xi_2)/k_BT} d\xi_1 d\xi_2 \tag{2.13}$$

Additional theoretical details and derivations may be found in Souaille and Roux [28]. Generic computer code for carrying out WHAM calculations is also provided in

this reference. The first PMFs computed for an iGluR LBD involved the GluA2 S1S2 construct [29]. The umbrella sampling simulations generating the PMFs employed a two-dimensional order parameter consisting of two distances between the lobes of the LBD (**Fig. 2-3**). The order parameter was inspired by the observation that the distances between T480 and S654, and between E402 and T686 each differ by ∼3.5 Å when comparing crystal structures of apo and agonist-bound LBDs [30]. Each pair of residues is hydrogen bonded in the agonist-bound structures but not in the apo structure. In the order parameter (ξ_1, ξ_2), ξ_1 is the distance between the centers of mass (COM) of residues 479-481 in lobe 1 and residues 654-655 in lobe 2, and ξ_2 is the distance between the COM of residues 401-403 in lobe 1 and residues 686-687 in lobe 2. These residues correspond to regions that are not in flexible loops. All non-hydrogen atoms were included in the selections. The rationale for using multiple residues in the definitions of ξ_1 and ξ_2 was to increase the likelihood that changes in the COM distances would represent large-scale conformational changes rather than fluctuations in individual residues. We should note that we initially attempted to define an angle to monitor opening and closing of the LBD, in a fashion similar to how the crystal structures were characterized. However, we found this choice of order parameter too unwieldy for use with the umbrella sampling restraining potentials. Other order parameters have been used by other investigators, for example, the distance between the Cα atoms of G451 and S652 [31]. Our chosen order parameter was intended to sufficiently capture the principal conformational transition inferred from LBD crystal structures. We acknowledge, however, that an order parameter with higher dimensionality would be required to capture motions involving additional degrees of freedom.

In our previous studies [29, 32, 33], we have found that umbrella sampling windows spaced 1 Å apart along ξ_1 and ξ_2 provide sufficient sampling overlap between neighboring windows. Windows spanning (ξ_1, ξ_2) = (8, 6) to (18,17) (in Å) appear to provide reasonable coverage of conformational space (**Fig. 2-4**). Initially, for GluA2, ∼200 windows were used [29]. Many of these windows, however, turned out to reside in regions of high conformational free energy. Subsequently, for GluN1, GluN2A,

and GluN3A, the number of windows was cut to the ∼100 that contribute most to $\langle\rho(\xi_1, \xi_2)\rangle$ [33]. Windows in the "corner" regions, which correspond to either large ξ_1 and small ξ_2 or small ξ_1 and large ξ_2, are the windows that are generally very high in conformational free energy and can be safely omitted from the computation.

2.3.4 Setting Up the System

Molecular simulations require complete protein models; that is, any missing non-hydrogen side chain or backbone atoms in a PDB entry must be built into the model. For building missing sidechains, we routinely use the program SCWRL [34]. For building missing backbone atoms, we use the loop-modeling routine of the program MODELER [35], also available via the ModLoop server (modbase.compbio.ucsf.edu). The next task is to generate initial protein configurations for each of the umbrella sampling windows. We use a targeted (biased potential) MD procedure to generate these LBD configurations. In this procedure, the Cα atoms of lobes 1 and 2 are separately restrained to their crystal structure configuration using an RMSD restraining potential while the lobes are pushed apart by performing dynamics in the presence of biasing potentials applied separately to ξ_1 and ξ_2. These simulations are run in vacuum. The intent of the RMSD restraints is to prevent the lobes from partially unfolding while ξ_1 and ξ_2 are rapidly moved apart. First, ξ_1 and ξ_2 are simultaneously incremented starting from a closed-conformation crystal structure. Next, ξ_1 and ξ_2 are individually incremented starting from configurations generated in the previous step. System setup concludes with solvating each conformer with explicit water and 100-150 mM NaCl. Nonzero total charge on the protein (and ligand, if present) may be neutralized by using excess Na^+ or Cl^- ions. We retain the positions of crystallographically-determined water molecules within the binding cleft.

2.3.5 Small Molecule Parametrization

For iGluR ligands (small organic molecules) that are not amino acids, the molecular mechanics parameters and charges will need to be determined. This can be accomplished a number of ways. One approach is to use the general AMBER force field

(GAFF) [36] together with Antechamber, a toolkit for automatic atom typing [37]. For CHARMM-consistent parameters, the CHARMM general force field (CGenFF) program can be used for automatic atom typing and assignment of parameters and charges by analogy [38,39]. Alternatively, a method called general automated atomic model parameterization (GAAMP) can be used to generate parameters based on the results of *ab initio* quantum mechanical (QM) calculations [40]. GAAMP can also be used to optimize existing parameter sets. Two web portals are available for automatic parameterization of small compounds using CGenFF: ParamChem (www.paramchem.org) and MATCH (brooks.chem.lsa.umich.edu/index.php?matchserver=submit). A web portal for GAAMP is at gaamp.lcrc.anl.gov. Finally, parameterization can be performed manually. Our approach for DNQX [29] included the following steps: (1) charge fit a geometry-optimized molecule to QM electrostatic potential maps, (2) optimize force constants by reproducing vibrational frequencies and potential energy distributions from QM calculations, and (3) optimize dihedral angle parameters by reproducing dihedral potential energy surfaces from QM calculations.

2.3.6 Equilibriation and Production

Each system must be equilibrated before sampling data are collected. The steps we routinely take are described below:

- First, the protein and ligand atoms are held fixed while dynamics are carried out on the solvent.
- Next, dynamics are carried out in the presence of stabilizing restraining potentials on the protein and ligand. These restraints are gradually reduced over the course of the equilibration, which is carried out in the constant atom number, volume, and temperature (NVT) ensemble.
- Finally, after these restraints have been removed, the simulation is transitioned to a constant pressure (NPT) ensemble.

In all steps of the equilibration involving non-fixed protein and ligand atoms, the umbrella restraints on ξ_1 and ξ_2 are also applied, as well as weak restraints on the COM of three regions in the core of lobe 1. The latter restraint, which does not

hinder inter-lobe dynamics, prevents translation and rotation of the protein during the production phase of the simulations.

The most effective way to implement the restraints described above vary depending on which MD simulation package is used. For example, in CHARMM, we use the restraint invoked by the keywords "CONS HARM" for the stabilizing potential. The "MMFP GEO SPHERE RCM DISTANCE" restraint is used for ξ_1 and ξ_2, and the "CONS HMCM" restraint is used for translation/rotation. The "MMFP" module allows the instantaneous values of ξ_1 and ξ_2 to be written on-the-fly. Details specifying how to use these keywords are included in CHARMM's "doc" (documentation) directory.

The simulations enter production phase after equilibration has completed. Here, the values for ξ_1 and ξ_2 in each window i are recorded, which contribute to the biased distribution function $\langle \rho(\xi_1, \xi_2) \rangle_{(i)}$ (see 2.5). It is critical that all distributions are carefully monitored to make sure there is sufficient overlap between neighboring windows, which is a requirement of WHAM [24]. If deficiencies are observed, additional windows at intermediate positions may be added. Additional windows may also be added to expand the boundaries of (ξ_1, ξ_2) if, after computing $W(\xi_1, \xi_2)$, it is determined that the free energy basins have not been sufficiently mapped. Automated adaptive umbrella sampling approaches may also be useful for optimizing the task of selecting windows [41].

2.3.7 Error Analysis

Once $W(\xi_1, \xi_2)$ has been computed, the quality of the sampling can be assessed. Two methods of analysis are block averaging and bootstrapping. In block averaging, as the name suggests, the sampling data (e.g., (ξ_1, ξ_2)) are divided into M segments, or "blocks", corresponding to a given simulation time. The PMF is then calculated for each block, and error bars can be generated to represent the deviation observed in the M PMFs. For a recent example of block averaging applied to umbrella sampling, see Park et al. [42]. In bootstrapping [43], "resampled" data sets are constructed by randomly selecting data points from the original sampling data. The resampling takes

place using time intervals estimated to correspond to the time it takes for correlated fluctuations in (ξ_1, ξ_2) to decay to zero. Previously, we used time intervals of 300 ps for GluN LBDs and 200 ps for the GluA2 LBD [33]. Resampling is repeated until each resampled data set is the same size as the original. The PMF is then calculated for each resampled set, and error bars can be generated as described above.

Ensuring that convergence in the sampling has been reached, using, for example, one or both of the methods described above, is an important step in molecular simulation studies. It is also a prerequisite for reasonably comparing the computed free energies with experimental measurements. For the computations described in this chapter, which are focused on large-scale conformational transitions, experimental measurements that, at least in principle, can be compared fairly directly include those that characterize conformational ensembles in solution, such as small-angle X-ray scattering (SAXS) [29, 44], and nuclear magnetic resonance (NMR) [45] and fluorescence resonance energy transfer (FRET) spectroscopy [46]. The computed free energy landscapes provide Boltzmann weights to be applied to the molecular configurations extracted from the simulations, which in turn provide a framework, based on molecular ensembles, for interpreting the measured data.

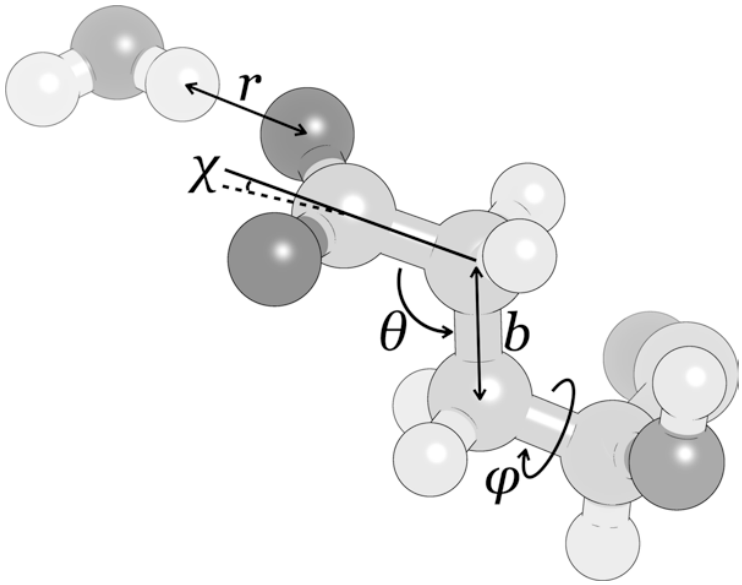

Fig. 2-1: **A molecular mechanics potential energy function** typically includes terms describing bond stretching (b), bond angle bending (θ), bond twsiting (ϕ), out-of-plane distortions (χ, and non-bonded interactions (r).

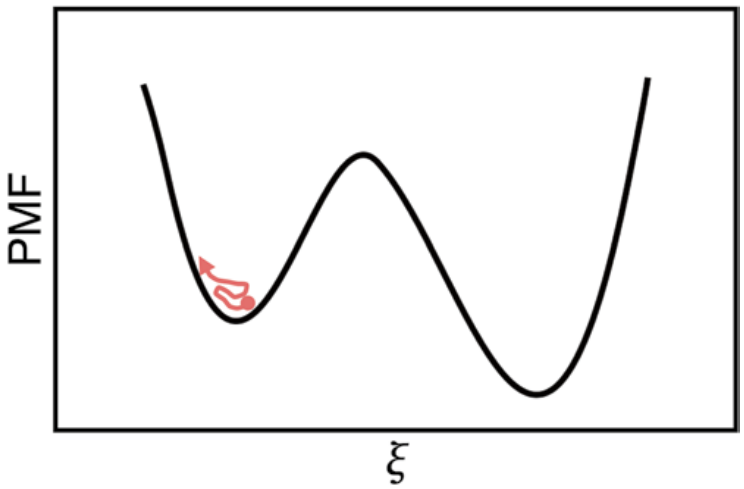

Fig. 2-2: **Free energy barriers.** In unbiased molecular simulations, one or more large free energy barries along an order parameter, ξ, can impede conformational sampling (depicted in *red*).

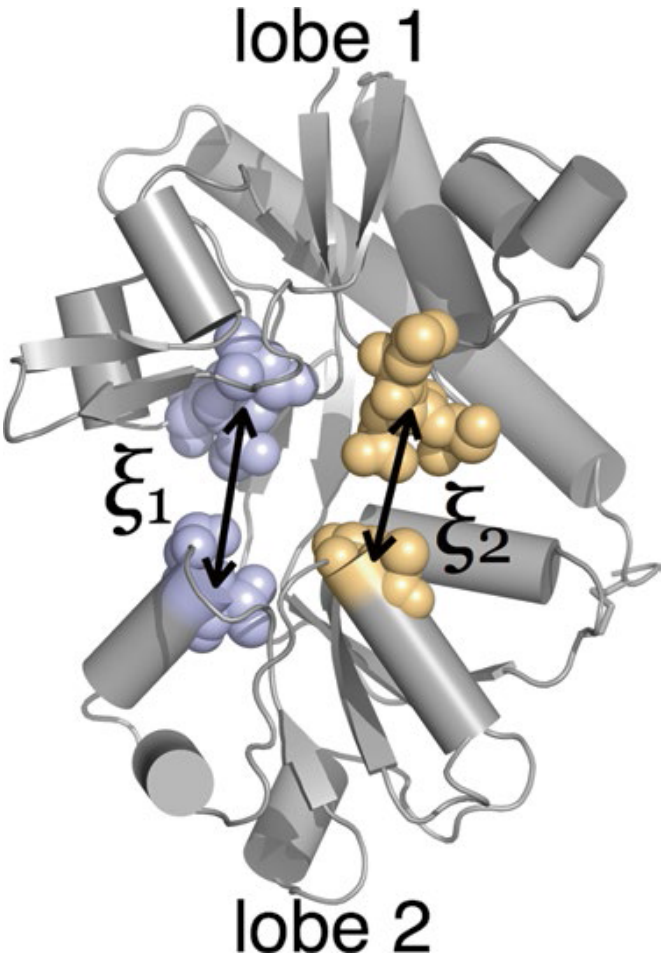

Fig. 2-3: **Order parameters.** Our chosen order parameter, (ξ_1, ξ_2), used to describe large-scale conformational transitions in the GluA2 LBD. ξ_1 and ξ_2 indicate the distances between the centers of mass of the atoms that are shown as spheres. In this figure, (ξ_1, ξ_2) = (12.8 Å, 11.4 Å)

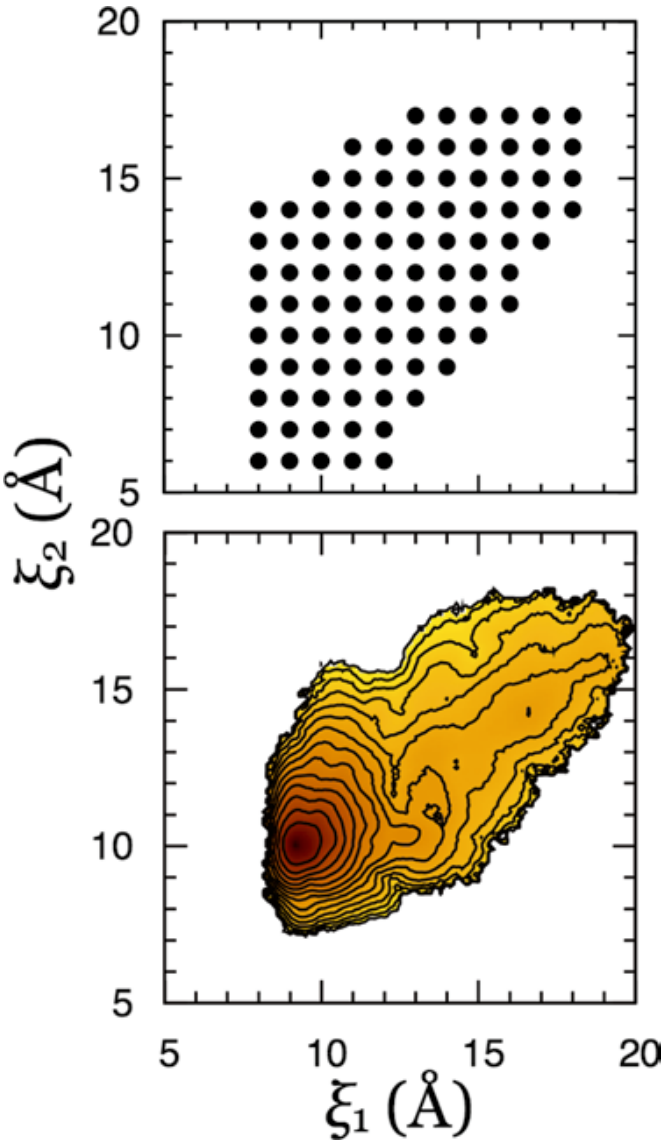

Fig. 2-4: **Umbrella sampling** involving the GluN1 glycine-bound LBD. The top panel shows the umbrella restraining potential equilibrium positions of the 93 windows that produce the free energy landscape shown in the bottom panel. In the bottom panel, each contour level corresponds to a difference of 1 kcal/mol, and darker regions are lower in free energy. See Yao et al. [33] for details.

References

[1] Kumar, J. and Mayer, M. L. Functional Insights from Glutamate Receptor Ion Channel Structures. *Annual review of physiology*, 75:313–337, 2013.

[2] Arinaminpathy, Y., Sansom, M. S. P., and Biggin, P. C. Binding site flexibility: molecular simulation of partial and full agonists within a glutamate receptor. *Molecular Pharmacology*, 69(1):11–18, January 2006.

[3] Mamonova, T., Speranskiy, K., and Kurnikova, M. Interplay between structural rigidity and electrostatic interactions in the ligand binding domain of GluR2. *Proteins: Structure, Function, and Bioinformatics*, 73(3):656–671, November 2008.

[4] Vijayan, R., Sahai, M. A., Czajkowski, T., and Biggin, P. C. A comparative analysis of the role of water in the binding pockets of ionotropic glutamate receptors. *Physical Chemistry Chemical Physics*, 12(42):14057–14066, October 2010.

[5] Cheng, X. and Ivanov, I. Molecular Dynamics. In *Computational Toxicology*, Methods in Molecular Biology, pages 243–285. Humana Press, Totowa, NJ, 2012. DOI: 10.1007/978-1-62703-050-2_11.

[6] Brooks, B., Brooks, C., MacKerell, A., Nilsson, L., Petrella, R., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R., Post, C., Pu, J., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D., and Karplus, M. CHARMM: The Biomolecular Simulation Program. *Journal of computational chemistry*, 30(10):1545–1614, July 2009.

[7] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic

Molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, May 1995.

[8] Scott, W. R. P., Hunenberger, P. H., Tironi, I. G., Mark, A. E., Billeter, S. R., Fennen, J., Torda, A. E., Huber, T., Kruger, P., and van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *The Journal of Physical Chemistry A*, 103(19):3596–3607, May 1999.

[9] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802, December 2005.

[10] Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, 26(16):1701–1718, December 2005.

[11] Bowers, K. J., Chow, E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Moraes, M. A., Sacerdoti, F. D., Salmon, J. K., Shan, Y., and Shaw, D. E. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, SC '06, New York, NY, USA, 2006. ACM.

[12] Guvench, O. and MacKerell, A. D. Comparison of Protein Force Fields for Molecular Dynamics Simulations. In *Molecular Modeling of Proteins*, Methods Molecular Biology, pages 63–88. Humana Press, 2008. DOI: 10.1007/978-1-59745-177-2_4.

[13] Humphrey, W., Dalke, A., and Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1):33–38, February 1996.

[14] Mackerell, A. D. Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry*, 25(13):1584–1604, October 2004.

[15] Jorgensen, W. L. and Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, March 1988.

[16] MacKerell, A. D., Feig, M., and Brooks, C. L. Improved Treatment of the Protein Backbone in Empirical Force Fields. *Journal of the American Chemical Society*, 126(3):698–699, January 2004.

[17] Perilla, J. R. and Woolf, T. B. Towards the prediction of order parameters from molecular dynamics simulations in proteins. *The Journal of Chemical Physics*, 136(16), April 2012.

[18] Alexandrov, V., Lehnert, U., Echols, N., Milburn, D., Engelman, D., and Gerstein, M. Normal modes for predicting protein motions: A comprehensive database assessment and associated Web tool. *Protein Science : A Publication of the Protein Society*, 14(3):633–643, March 2005.

[19] Chennubhotla, C. and Bahar, I. Signal propagation in proteins and relation to equilibrium fluctuations. *PLoS computational biology*, 3(9):1716–1726, September 2007.

[20] Skjaerven, L., Hollup, S. M., and Reuter, N. Normal mode analysis for proteins. *Journal of Molecular Structure: THEOCHEM*, 898(1):42–48, March 2009.

[21] Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *The Journal of Chemical Physics*, 3(5):300–313, May 1935.

[22] Roux, B. The calculation of the potential of mean force using computer simulations. *Computer Physics Communications*, 91(1):275–282, September 1995.

[23] Torrie, G. M. and Valleau, J. P. Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chemical Physics Letters*, 28(4):578–581, October 1974.

[24] Kastner, J. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(6):932–942, November 2011.

[25] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, 13(8):1011–1021, October 1992.

[26] Ferrenberg, A. M. and Swendsen, R. H. Optimized Monte Carlo data analysis. *Physical Review Letters*, 63(12):1195–1198, September 1989.

[27] Boczko, E. M. and Brooks, C. L. Constant-temperature free energy surfaces for physical and chemical processes. *The Journal of Physical Chemistry*, 97(17):4509–4513, April 1993.

[28] Souaille, M. and Roux, B. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer Physics Communications*, 135(1):40–57, March 2001.

[29] Lau, A. Y. and Roux, B. The Free Energy Landscapes Governing Conformational Changes in a Glutamate Receptor Ligand-Binding Domain. *Structure (London, England : 1993)*, 15(10):1203–1214, October 2007.

[30] Armstrong, N. and Gouaux, E. Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron*, 28(1):165–181, October 2000.

[31] Mamonova, T., Yonkunas, M. J., and Kurnikova, M. G. Energetics of the cleft closing transition and the role of electrostatic interactions in conformational rearrangements of the glutamate receptor ligand binding domain. *Biochemistry*, 47(42):11077–11085, October 2008.

[32] Lau, A. Y. and Roux, B. The hidden energetics of ligand binding and activation in a glutamate receptor. *Nature Structural & Molecular Biology*, 18(3):283–287, March 2011.

[33] Yao, Y., Belcher, J., Berger, A. J., Mayer, M. L., and Lau, A. Y. Conformational analysis of NMDA receptor GluN1, GluN2, and GluN3 ligand-binding domains reveals subtype-specific characteristics. *Structure (London, England: 1993)*, 21(10):1788–1799, October 2013.

[34] Krivov, G. G., Shapovalov, M. V., and Dunbrack, R. L. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins*, 77(4):778–795, December 2009.

[35] Fiser, A. and Sali, A. ModLoop: automated modeling of loops in protein structures. *Bioinformatics (Oxford, England)*, 19(18):2500–2501, December 2003.

[36] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174, July 2004.

[37] Wang, J., Wang, W., Kollman, P. A., and Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling*, 25(2):247–260, October 2006.

[38] Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. 52.

[39] Vanommeslaeghe, K., Raman, E. P., and MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of bonded parameters and partial atomic charges. *Journal of chemical information and modeling*, 52(12):3155–3168, December 2012.

[40] Huang, L. and Roux, B. Automated Force Field Parameterization for Nonpolarizable and Polarizable Atomic Models Based on Ab Initio Target Data. *Journal of Chemical Theory and Computation*, 9(8):3543–3556, August 2013.

[41] Wojtas-Niziurski, W., Meng, Y., Roux, B., and Berneche, S. Self-Learning Adaptive Umbrella Sampling Method for the Determination of Free Energy Land-

scapes in Multiple Dimensions. *Journal of chemical theory and computation*, 9(4):1885–1895, April 2013.

[42] Park, S., Kim, T., and Im, W. Transmembrane Helix Assembly by Window Exchange Umbrella Sampling. *Physical review letters*, 108(10):108102, March 2012.

[43] Efron, B. and Tibshirani, R. J. *An Introduction to the Bootstrap*. CRC Press, May 1994.

[44] Madden, D. R., Armstrong, N., Svergun, D., Perez, J., and Vachette, P. Solution X-ray Scattering Evidence for Agonist- and Antagonist-induced Modulation of Cleft Closure in a Glutamate Receptor Ligand-binding Domain. *Journal of Biological Chemistry*, 280(25):23637–23642, June 2005.

[45] Ahmed, A. H., Loh, A. P., Jane, D. E., and Oswald, R. E. Dynamics of the S1s2 Glutamate Binding Domain of GluR2 Measured Using 19f NMR Spectroscopy. *Journal of Biological Chemistry*, 282(17):12773–12784, April 2007.

[46] Landes, C. F., Rambhadran, A., Taylor, J. N., Salatan, F., and Jayaraman, V. Structural landscape of the isolated ligand binding domain of single AMPA receptors. *Nature chemical biology*, 7(3):168–173, March 2011.

Chapter 3 - Molecular lock regulates binding of glycine to a primitive NMDA receptor

Reproduced in part from:

Proceedings of the National Academy of Sciences USA, 113(44) E6786-E6795, 2016

DOI:10.1073/pnas.1607010113

by

Alvin Yu, Robert Alberstein, Alecia Thomas, Austin Zimmet, Richard Grey, Mark L. Mayer, and Albert Y. Lau

Glycine-activated ionotropic glutamate receptors (iGluRs) encoded in ctenophore genomes are evolutionary precursors to NMDA receptors, which play important roles in synaptic plasticity. Ctenophore iGluRs feature a distinct interdomain salt bridge in the ligand-binding domain, a molecular lock, thus far not found in iGluRs of other organisms. We use a combination of crystallographic, biochemical, electrophysiological, and computational approaches to elucidate the role of this molecular lock in a ctenophore iGluR. We find that perturbations to the lock can tune receptor kinetics and thermodynamics over very broad ranges. We also find that the strategic location of the lock may be the basis for the ligand-binding domain's extraordinarily high affinity for glycine.

The earliest metazoan ancestors of humans include the ctenophore *Mnemiopsis leidyi*. The genome of this comb jelly encodes homologs of vertebrate ionotropic glutamate receptors (iGluRs) that are distantly related to glycine-activated NMDA receptors and that bind glycine with unusually high affinity. Using ligand-binding domain (LBD) mutants for electrophysiological analysis, we demonstrate that perturbing a ctenophore-specific interdomain Arg-Glu salt bridge that is notably absent from vertebrate AMPA, kainate, and NMDA iGluRs greatly increases the rate of recovery from desensitization, while biochemical analysis reveals a large decrease in affinity for glycine. X-ray crystallographic analysis details rearrangements in the binding pocket stemming from the mutations, and molecular dynamics simulations suggest that the interdomain salt bridge acts as a steric barrier regulating ligand binding and that the free energy required to access open conformations in the glycine-bound LBD is largely responsible for differences in ligand affinity among the LBD variants.

3.1 Background

Glutamate receptor ion channels (iGluRs) are membrane proteins that mediate excitatory synaptic transmission in the brain by detecting release of the amino acid glutamate from nerve terminals [1]. In combination with GluN2 subunits, which bind glutamate, NMDA subtype iGluRs use glycine as a coagonist, which binds to GluN1, GluN3A, and GluN3B subunits [2–6]. NMDA receptors play key roles in synaptic

plasticity and memory formation, and mutations of NMDA receptor genes underlie a diverse set of neurological and psychiatric diseases [7]. Like all iGluRs, NMDA receptors are assembled from modular subunits containing amino terminal and S1S2 ligand binding domains (LBDs), which can be genetically isolated and expressed as soluble proteins for biochemical and structural analysis [4, 8–10]. The LBDs of both the glutamate and glycine binding subunits are clamshell-shaped proteins of molecular mass around 30 kDa in which two lobes are connected by a hinge formed by antiparallel β-strands; in the activated state, ligands are trapped in a cavity formed when the clamshell closes. Strikingly, the volume of the ligand binding cavity for the GluN1, GluN3A, and GluN3B subunits is just large enough to accommodate glycine, whereas iGluR glutamate binding subunits have cavities that are four to five times larger and bind both glutamate and up to six or seven water molecules [4, 10–13].

We recently reported the discovery of glycine-activated iGluRs from the comb jelly *Mnemiopsis leidyi* and the sea gooseberry *Pleurobrachia bachei*, candidates for earliest lineage metazoans, for which ML032222a and PbGluR3 glycine complex crystal structures reveal a salt bridge at the perimeter of the ligand binding cleft [14]. This salt bridge links the upper and lower lobes of the LBD in the closed cleft glycine-bound conformation. Ctenophore iGluR subunits bind glycine with such high affinity that the ligand cannot be removed by exhaustive dialysis, suggesting an unusually stable ligand-bound closed-cleft conformation, perhaps stabilized by the interdomain salt bridge. Prior electrophysiological and crystallographic studies on vertebrate AMPA and kainate subtype iGluRs revealed that the stability of the closed cleft conformation is determined not only by contacts of the LBD with the neurotransmitter ligand but also by contacts formed between the upper and lower lobes of the clamshell assembly that occur only in the ligand-bound closed-cleft conformation [15, 16]. Comparison of crystal structures of ctenophore iGluR LBDs with those of vertebrate NMDA receptor GluN1 and GluN3 subunits that also bind glycine, but for which apo proteins can be prepared without difficulty [4, 10], reveals that the salt bridge is unique to ctenophore iGluRs, further suggesting that it might underlie the high stability of the glycine complex.

To investigate this, we prepared ML032222a mutant proteins and analyzed their ligand binding properties using electrophysiological, biochemical, and crystallographic techniques. To gain further insight into how these mutants perturb large-scale LBD dynamics, we computed conformational free energy landscapes for the apo state and glycine complexes of wild-type (WT) ML032222a and the R703K and E423S mutants, which weaken and break the interdomain salt bridge, respectively. This analysis reveals that, similar to vertebrate GluN1 and GluN3 glycine binding subunits, the apo state for ML032222a can access closed cleft conformations, although it is more stable in slightly open conformations. The R703K and E423S mutants destabilize closed cleft conformations for the glycine complex. Conformational dynamics inferred from the free energy landscapes suggest that the interdomain salt bridge is positioned at the most likely point of ligand entry to (and exit from) the binding pocket and thus acts as a steric barrier regulating the binding and dissociation of glycine

3.2 Results

3.2.1 Interdomain Salt Bridge Mutants Regulate Recovery from Desensitization.

ML032222a from the comb jelly *M. leidyi* forms a homomeric glycine-activated ion channel with rapid onset but extremely slow recovery from desensitization [14]. We hypothesized that the unusually slow recovery from desensitization for ML032222a, which lasts for several minutes, occurs because the glycine bound LBD is trapped in a closed cleft conformation by a lock at the entrance to the binding site (**Fig. 3-1**A) and that the receptor cannot recover from desensitization until the clamshell opens and glycine dissociates. In the ML032222a LBD crystal structure, this lock is formed by an interdomain salt bridge between Glu423 in lobe 1 and Arg703 in lobe 2 (**Fig. 3-1**B). To weaken this interaction we prepared two mutants, ML032222a R703K and ML032222a E423S, and expressed them in *Xenopus* oocytes for electrophysiological analysis. Twin pulse applications of glycine for WT ML032222a revealed that recovery from desensitization required several minutes (**Fig. 3-1**C), with progressively faster

recovery for the R703K (**Fig. 3-1**D) and E423S mutants (**Fig. 3-1**E); when fit with exponential functions, the rate constant for recovery from desensitization increased 4-fold for the R703K mutant and 61-fold for the E423S mutant (**Fig. 3-1**F) [rate constant: 0.016 ± 0.003 s^{-1} for WT ($n = 4$), 0.061 ± 0.003 s^{-1} for R703K ($n = 3$), and 0.96 ± 0.06 s^{-1} for E423S ($n = 3$)].

3.2.2 Interdomain Salt Bridge Mutants Lower Affinity for Glycine.

The S1S2 LBD of ML032222a expressed as a soluble protein binds glycine with an affinity of 2.3 nM and is resistant to proteolysis by trypsin unless it is dialyzed in the unfolded state to remove endogenous glycine that remains bound during purification [14]. Proteolysis protection assays revealed that similar to the WT protein, unfolding and dialysis in 4 M guanadinium was also required to prepare apo protein for the R703K mutant, whereas by contrast, for the E423S mutant, apo protein could be prepared by exhaustive dialysis without the need for unfolding (**Fig. 3-2**A). Similar to WT ML032222a, both mutants became resistant to digestion by trypsin following addition of glycine to the refolded apo protein, indicating that they remain competent to bind glycine (**Fig. 3-2**A), consistent with the results of electrophysiological experiments that show activation of ion channel gating by glycine (**Fig. 3-1**). Using refolded or dialyzed proteins as appropriate, we then performed isothermal titration calorimetry (ITC) to measure the thermodynamics of ligand binding. In previously published work, we obtained a K_d of 2.7 nM (95% CI: 1.03, 4.95 nM) for WT ML032222a; binding was strongly exothermic (ΔG, -11.50 kcal/mol; ΔH, -17.34 kcal/mol; TΔS, -5.84 kcal/mol) [14]. For the R703K mutant, the K_d increased 10-fold to 28.3 nM (95% CI: 19.1, 40.9 nM); binding remained exothermic (ΔG, -10.12 kcal/mol; ΔH, -21.13 kcal/mol; TΔS, -11.01 kcal/mol) (**Fig. 3-2**B). For the E423S mutant, the K_d increased nearly three orders of magnitude to 2.48 μM (95% CI: 2.16, 2.84 μM); binding remained exothermic (ΔG, -7.52 kcal/mol; ΔH, -18.11 kcal/mol; TΔS, -10.59 kcal/mol) (**Fig. 3-2**C). These results suggest that faster recovery from desensitization for the mutant proteins observed in electrophysiological experiments

(**Fig. 3-1**) is a direct consequence of disrupting the interdomain salt bridge lock.

3.2.3 Crystal Structures of Interdomain Salt Bridge Mutants.

In the presence of glycine, the R703K and E423S mutant S1S2 LBDs crystallized as dimer assemblies in the same space group as the WT protein and diffracted X-rays to Bragg spacings of 1.34 and 1.28 Å, respectively (**Table 3.1**); electron density was well resolved for the mutant side chains, the glycine ligand, and the numerous solvent molecules flanking the binding site (**Fig. 3-3**). Least squares superpositions using 203 Cα atom coordinates for the R703K and E423S mutants revealed essentially identical conformations to the WT protein (RMSD, 0.26 and 0.20 Å, respectively). In all three structures, glycine binds in a cavity, the sides and top of which are formed by the apposed surfaces of lobes 1 and 2. In the R703K and E423S mutants the bound glycine ligand forms the same series of contacts as occurs in the WT protein; the glycine ligand α-carboxyl group is bound by the main chain amide of Ser499 and the side chain of Arg504 in lobe 1 and by the main chain amide nitrogen of His505 in lobe 2; the glycine α-amino group is bound by the main chain carbonyl oxygen of Asp497 and the side chain hydroxyl group of Ser499 in lobe 1 and by the side chain of Glu747 in lobe 2 (**Fig. 3-3**).

Calculation of solvent accessible cavity volumes, with a probe radius of 1.4 Å, gave similar values for WT ML032222a and for the R703K and E423S mutants (58.8 $\pm$ 0.01, 56.6 $\pm$ 0.1, and 59.0 $\pm$ 0.03 Å^3, respectively; mean $\pm$ SEM; $n = 3$); however, close inspection of the cavity volumes revealed subtle changes in shape due to movements resulting from introduction of the mutant side chains (**Fig. 3-4**). As a result, in lobe 2, the binding site surfaces for the mutant structures are distinct from those for WT ML032222a (**Fig. 3-3**A). In the R703K mutant, the side chain of Glu423 rotates 43° around Chi2 and 31° around Chi3 such that the Glu OE2 carboxyl group oxygen atom moves 1.8 Å away from the glycine ligand to form an ion pair contact with the R703K mutant side chain, breaking the carboxylate-carboxylate contact with Asp497 found in the WT structure (**Fig. 3-3**B). This movement creates a bulge occupied by a water molecule (W9), which is coordinated by the Glu423, Asp497, and Glu747

side chains (**Fig. 3-3**B). In the E423S mutant (**Fig. 3-3**C), a pair of water molecules equivalent to W7 and W8 in the WT structure moves down toward the glycine ligand such that W7 is positioned exactly at the location of the Glu423 OE1 carboxyl group oxygen atom in the WT protein and makes hydrogen bond contacts with the mutant Ser423, Asp497, and Arg703 side chains (**Fig. 3-3**C).

3.2.4 Conformational Dynamics of the LBD.

To gain further insight into the structural thermodynamics of the ML032222a LBD, we computed the conformational free energy landscapes, or potentials of mean force (PMFs), of the WT S1S2 LBD and the R703K and E423S mutants for both apo- and glycine-bound forms. The PMFs were computed using umbrella sampling molecular dynamics (MD) simulations in which a 2D order parameter (ξ_1, ξ_2) was used to characterize large-scale conformational changes in the LBD, where ξ_1 and ξ_2 are center of mass distances between atom selections in lobes 1 and 2 (**Fig. 3-5**A). Umbrella sampling using an analogous order parameter was used previously to characterize conformational transitions in AMPA and NMDA receptor LBDs, and the results were found to be consistent with small-angle X-ray scattering (SAXS), ligand-binding affinity, and single-molecule FRET (smFRET) studies [17–21]. Because in the crystal structures the side chains of Glu423 and Asp497 are in close proximity, pKa calculations were carried out using PROPKA [22] to determine which protonation states of these two residues to use in our simulations. On the basis of these calculations, Glu423 was not protonated in the WT and R703K simulations, whereas Asp497 was protonated. For the E423S LBD, the pKa calculations suggested that Asp497 should not be protonated in the glycine complex but should be protonated in the apo state. These protonation states were therefore used in the E423S simulations.

The WT apo LBD PMF features a broad free energy basin, indicating conformational flexibility in the absence of ligands (**Fig. 3-5**B). The global free energy minimum is located at (ξ_1, ξ_2) = (10.7, 13.3 Å). As expected, the WT glycine-bound LBD PMF features an energy basin that is narrower, indicating stabilized closed cleft conformations. The global free energy minimum is located at (ξ_1, ξ_2) = (9.0, 11.3

Å). In the glycine complex crystal structure, the LBDs in the dimer assembly have nearly identical conformations, which are located nearby at (ξ_1, ξ_2) = (8.7, 10.7 Å) and (8.7, 10.8 Å) for chains A and B, respectively. Although there is some overlap in the free energy basins between the apo- and glycine-bound LBD PMFs, these LBDs populate distinct conformational ensembles (**Fig. 3-6**). Within 1 kcal/mol of the minimum, the WT, apo LBD is able to access more open conformations, ranging from (ξ_1, ξ_2) = (9.8, 11 Å) to (ξ_1, ξ_2) = (11.2, 14.2 Å), whereas the WT, glycine-bound LBD occupies more closed conformations, from (ξ_1, ξ_2) = (8.6, 10.3 Å) to (ξ_1, ξ_2) = (9.6, 11.1 Å). Interestingly, the apo ML032222a LBD exhibits a single conformational free energy minimum, which is observed in glutamate-binding LBDs from the iGluRs GluA2 and GluN2A rather than multiple conformational free energy minima, as seen in glycine-binding LBDs from the iGluRs GluN1 and GluN3A [19].

For both the R703K and E423S mutants, the apo LBD PMFs exhibit slightly narrowed free energy basins compared with the WT protein (**Fig. 3-5** C and D), although all three PMFs are fairly similar. The glycine-bound LBD PMFs, however, exhibit shallower free energy funnels, indicating that open LBD conformations are easier to access. For apo R703K, the global free energy minimum is located at (ξ_1, ξ_2) = (10.5, 12.7 Å), and for apo E423S, it is located at (ξ_1, ξ_2) = (10.1, 12.0 Å). For glycine-bound R703K, the global free energy minimum is located at (ξ_1, ξ_2) = (9.4, 11.3 Å), and for glycine-bound E423S, it is located at (ξ_1, ξ_2) = (9.3, 12.2 Å). Perhaps due to crystal lattice packing effects, the structures of the R703K and E423S mutant glycine complexes adopt slightly more closed conformations, (ξ_1, ξ_2) = (8.8, 10.4 Å) and (8.7, 10.6 Å), and (ξ_1, ξ_2) = (8.7, 10.6 Å) and (8.7, 10.7 Å), for chains A and B, respectively. **Fig. 3-7** shows the SDs for the 2D PMFs, as determined by block averaging.

3.2.5 Stability of the Interdomain Salt Bridge.

How stable is the salt bridge formed by Glu423 with the Arg or Lys residue at position 703? To examine the structural stability of this salt bridge in the WT and R703K mutant LBDs as a function of LBD conformation in our simulations, we plotted side

chain distances between residues 703 and 423 vs. (ξ_1, ξ_2) (**Fig. 3-8** A and B and **Fig. 3-9** A and B). The statistics for the WT glycine complex and apo systems were calculated using ∼180,000 observations (snapshots extracted from the simulations) for each state. The statistics for each of the R703K systems were calculated using ∼190,000 observations. Average distances (and their SDs) were calculated between the closest nitrogen and oxygen atoms in the side chains of either Arg703 or Lys703 and Glu423, respectively. The distribution of average distances is not dramatically different between the apo LBD and the glycine complex for either the WT or R703K LBD. Residues 703 and 423, however, tend to stay in closer proximity for the WT LBD than for the R703K variant, reflecting, unsurprisingly, a stronger interaction for an arginine-glutamate pair than for a lysine-glutamate pair. The SDs of the average distances, however, are greater in the WT LBD, especially for the glycine complex, than for the R703K LBD. These results indicate that the Arg703-Glu423 bridge is able to remain intact over a larger range of LBD conformations than the Lys703-Glu423 bridge. The regions of large fluctuations in the Arg703-Glu423 distance, i.e., the SDs, indicate LBD conformations for which this salt bridge is strained.

The rotation of the Glu423 side chain away from Asp497 observed in the crystal structure of the R703K glycine complex LBD is also seen in the simulations of both WT and R703K proteins. Interestingly, this rotation occurs much more in the apo LBDs than in the glycine complexes (**Fig. 3-8** C and D and **Fig. 3-9** C and D). Although the bound glycine ligand does not directly contact Glu423, the ligand may stabilize the binding pocket residues in such a way that disfavors rotation of Glu423 out of the pocket. It is also possible that the orientation of Glu423 could impact stability of the Arg/Lys703-Glu423 salt bridge. If so, this would imply that the salt bridge "locks" apo and glycine-bound LBDs to different extents. The similarity in the PMFs for the apo WT, R703K, and E423S LBDs suggests the salt bridge conformationally restricts the glycine-bound LBD more than the apo state.

3.2.6 Occupancy of Binding Pocket Waters W9 and W7.

Specific water molecules in the binding pocket of the crystal structures for the R703K and E423S LBDs, i.e., waters W9 and W7, respectively, have been described above. How stable is the positioning of these waters, and what role might they play in stabilizing these structures? To examine water stability as a function of LBD conformation in our simulations, we plotted average occupancies of the water coordination sites vs. (ξ_1, ξ_2) (**Fig. 3-10**). For the R703K LBD, W9 is coordinated by the side chains of Glu423, Asp497, and Glu747. For the E423S LBD, W7 is coordinated by the side chains of Ser423, Asp497, and Arg703. In our simulation trajectories, a coordination site is considered occupied if a water's oxygen atom is within 3.2 Å of a nitrogen or oxygen atom belonging to each of the three coordinating side chains. For the R703K LBD, W9 is coordinated with high probability in certain closed-cleft conformations, especially (ξ_1, ξ_2) $\sim$ (11.5, 10.5 Å) for the apo state and (ξ_1, ξ_2) $\sim$ (10.0, 9.0 Å) for the glycine complex. For the E423S LBD, W7 is coordinated with high probability in a subset of closed-cleft conformations for the glycine complex, but it is coordinated with very low probability in the apo state in any conformation. The water occupancies for the glycine complex simulations are consistent with the observation of bound waters W7 and W9 observed in the crystal structures (see the locations of the crystal structures in **Fig. 3-5** C and D). The computed water occupancies also reveal other LBD conformations that give rise to side chain arrangements favorable to the coordination of W9 and W7, suggesting an important role in the ligand-binding mechanism. The statistics for each of the R703K systems and the apo E423S system were calculated using $\sim$190,000 observations (snapshots extracted from the simulations); the statistics for the glycine-bound E423S system were calculated using $\sim$132,000 observations.

3.2.7 The Conformational Free Energy Associated with Ligand Binding.

The total free energy of ligand binding into an LBD can be broadly decomposed into free energy contributions resulting from ligand docking into the binding pocket and

contributions resulting from large-scale LBD conformational changes, i.e., domain closure [18]. To compute the LBD conformational free energy associated with glycine binding, the extent to which the LBD has to open to allow glycine entry (and exit) must first be determined. Access tunnels into the binding pocket were analyzed using CAVER Analyst [23] for a selection of LBD conformations extracted from our simulations (**Fig. 3-11**A). The residues forming the bottleneck for the passage of a glycine molecule, which was estimated to have a minimum van der Waals radius of 2.62 Å, are Phe469 in lobe 1 and Arg/Lys703 in lobe 2, with Glu/Ser423 (in lobe 1) also lining the passageway. The minimum distance between the side chains of residues 469 and 703 that allows glycine passage is ∼10 Å, as measured between Cγ atoms. To examine glycine accessibility as a function of LBD conformation, we plotted average distances (and their SDs) between residues 469 and 703 vs. (ξ_1, ξ_2) (**Fig. 3-11** B-D and **Fig. 3-12** A-C). Statistics for the WT glycine complex and apo systems were calculated using ∼180,000 observations (snapshots extracted from the simulations) for each state, ∼190,000 observations for the R703K systems, and ∼132,000 and ∼190,000 observations for the E423S glycine complex and apo systems, respectively.

The relative probability for the LBD to adopt open conformations (allows glycine access) vs. closed conformations (disallows glycine access) can be computed as follows:

$$\frac{\Omega_{\text{Open}}}{\Omega_{\text{Closed}}} = \frac{\int_{\text{Open}} \exp(-W(\xi_1, \xi_2/k_B T)) d\xi_1 d\xi_2}{\int_{\text{Closed}} \exp(-W(\xi_1, \xi_2/k_B T)) d\xi_1 d\xi_2} \tag{3.1}$$

where Ω is the partition function for the LBD's conformational state, Open indicates the region of integration in the PMF, written as $W(\xi_1, \xi_2)$, that corresponds to values of (ξ_1, ξ_2) in which the average distance between residues 469 and 703 $\geq$ 10 Å; Closed indicates all other regions in the PMF; k_B is Boltzmann's constant, and T is temperature. The equation above can be related to the free energy difference between the open and closed state ensembles as follows:

$$\Delta G_{\text{conf}} = G_{\text{Open}} - G_{\text{Closed}} = -k_B T \log \left[\frac{\Omega_{\text{Open}}}{\Omega_{\text{Closed}}} \right] \tag{3.2}$$

Finally, the difference in ΔG_{conf} between glycine-bound and apo LBDs is given by

$$\Delta\Delta G_{\text{conf}} = \Delta G_{\text{conf,Gly}} - \Delta G_{\text{conf,Apo}} \tag{3.3}$$

$\Delta\Delta G_{\text{conf}}$ represents the difference in conformational free energy between the glycine-bound and apo LBDs for adopting open vs. closed states. In other words, $\Delta\Delta G_{\text{conf}}$ is the conformational free energy gained by the system on glycine binding for accessing open states. It is important to note that $\Delta\Delta G_{\text{conf}}$ does not correspond to the total binding free energy because free energy contributions associated with, for example, translation of the ligand from bulk solvent into the binding site have not been taken into account in this calculation. $\Delta\Delta G_{\text{conf}}$ alone, therefore, cannot be directly compared with experimentally measured binding free energies. If the differences in glycine binding affinity among the WT, R703K, and E423S proteins are dominated by changes in LBD conformational free energies, however, then the trend in calculated $\Delta\Delta G_{\text{conf}}$ values should follow the measured affinities. ΔG_{conf} were computed for each PMF, using an integration region of $\{\xi_1 \geq 12$ Å and $\xi_2 \geq 12$ Å$\}$ for open conformations. The results are provided in **Table 3.2**. The trend $\Delta\Delta G_{\text{conf,WT}} > \Delta\Delta G_{\text{conf,R703K}} > \Delta\Delta G_{\text{conf,E423S}}$ is consistent with the measured affinities, although $(\Delta\Delta G_{\text{conf,WT}} - \Delta\Delta G_{\text{conf,R703K}})$ is larger than expected from the measured affinities and $(\Delta\Delta G_{\text{conf,R703K}} - \Delta\Delta G_{\text{conf,E423S}})$ is smaller. These discrepancies might be due to possible electrostatic interactions between the glycine ligand and residues 703 and/or 423, which were not examined in this study. Varying the region of integration by ±1 Å in either ξ_1 or ξ_2 does not change the trend or the $\Delta\Delta G$ values significantly. $\Delta G_{\text{conf,Apo}}$ is ∼3 kcal/mol for all three proteins, so $\Delta G_{\text{conf,Gly}}$ therefore appears to be driving the changes in glycine affinity for the LBD variants; the presence of glycine in the binding pocket makes it more difficult for the WT LBD to adopt open conformations compared with the mutants.

3.3 Discussion

Multiple glycine-activated iGluR subunits have recently been discovered by genome sequencing of ctenophore species [24], including the comb jelly *M. leidyi* [14, 25]. A special feature of these subunits is an interdomain salt bridge that has not been found in iGluR subunits encoded in the genomes of a wide range of other organisms, including mammals. Using a combination of crystallographic, biochemical, functional, and computational approaches, we probed the effects of mutations in the salt bridge to try to understand its role in ctenophore iGluRs. The mutants R703K and E423S, which progressively weaken the interaction between residues 423 and 703, which form the salt bridge in the WT protein, result in increased rates of recovery from desensitization. This observation is consistent with our biochemical measurements that demonstrate that the mutations decrease LBD affinity for glycine, thus establishing that the lifetime of the desensitized state is determined in part by stability of the agonist receptor complex. Crystal structures of the mutant LBDs show that rearrangements in the binding cleft are mostly limited to residues 703 and 423 and a surrounding network of water molecules. Computational studies of the interdomain salt bridge indicate that, in the WT LBD, the Arg703-Glu423 bridge is stable over a wider range of LBD conformations than the Lys703-Glu423 counterpart present in the R703K mutant. The crystallographic waters W9 and W7, which form part of a hydrogen bonding network in the binding pockets of the R703K and E423S glycine-complex LBDs, respectively, are also found to be stably coordinated in the simulations.

Prior studies on GluA2, GluK2, and the NMDA receptor GluN1 and GluN2 subunits revealed only modest tuning of ligand apparent affinity by direct lobe 1-lobe 2 interactions [15, 16, 26–29]. The 1,000-fold stabilization of glycine binding by the ML032222a salt bridge demonstrates that much stronger modulation is possible for the LBD clamshell assembly. Functionally, a very high affinity for glycine has severe consequences: ML032222a desensitizes rapidly and takes many minutes to recover before it can respond to glycine again. On the other hand, the glycine-binding subunits of vertebrate NMDA receptors do not function as conventional ligand-gated ion

channel subunits, which sense the pulsatile release of neurotransmitter at synapses, but instead are likely to be continually bound by glycine or D-serine [30, 31]. Our functional and biochemical analyses of perturbations to the salt bridge show how interdomain contacts can tune receptor kinetics and thermodynamics over a very broad range.

Based on the electrophysiological and biochemical data reported in this paper, as well as prior studies of ML032222a and an iGluR LBD from another ctenophore, the sea gooseberry *P. bachei* [14], we had expected the salt bridge to simply stabilize a closed conformation for an open/closed transition. It was unclear, however, whether this would be enough to confer the extraordinarily high stability observed for the glycine complex. The results of our simulations reveal a surprisingly more complex situation than we had previously envisioned. The conformational free energy analysis reveals that the energy required to open the apo LBD sufficiently to allow glycine binding is similar among the WT and mutant proteins. Opening the glycine-bound LBD enough to allow glycine to exit, however, is energetically more expensive for the WT protein than for the mutants. Because the salt bridge, in particular Arg703, is positioned at what appears to be the likely point of entry into the binding pocket, it serves as a barrier to incoming ligands. The salt bridge also blocks ligand egress from the pocket if the binding pathways are assumed to be reversible. The effects of the interdomain salt bridge may therefore be twofold: it restricts conformational flexibility of the glycine-bound LBD while impeding the passage of glycine in and out of the pocket. It is also possible that the strength of the salt bridge might be tuned by the rotameric state of Glu423. The rotameric state appears roughly bimodal with respect to whether the LBD is apo or glycine bound. Combined with the effects mentioned above, a stronger salt bridge for the liganded complex would contribute to the LBD's high affinity for glycine. At a cleft opening of $(\xi_1, \xi_2) = (13, 16$ Å), which is sufficient to allow entry of a glycine ligand, the side chains of Arg703 and Glu423 are separated by $\sim$8 Å. Is it possible that the glycine might transiently interact with the guanidine of Arg703 and/or the carboxylate of Glu423 during ligand entry or exit? Thus far, we have been unsuccessful in simulating glycine binding and unbinding,

using approaches free of biasing potentials, to directly address this question, but we speculate that the LBDs for other iGluR subtypes may be similarly complex in terms of strategically positioned barriers.

The emerging structural biology of iGluR LBD conformational ensembles reveals that, although agonist binding results in the free energy landscapes adopting a single-minimum topology, apo LBDs can exhibit either single minima (as seen for ML032222a, GluA2, and GluN2A) or multiple minima (as seen for GluN1 and GluN3A) [17–19]. Glutamine-binding protein (GlnBP) is an example of a periplasmic binding protein (PBP) that exhibits a single stable conformational state for the apo protein [32], whereas maltose-binding protein (MBP) is an example of a PBP that exhibits metastable states for both open and closed conformations for the apo protein [33, 34]. It should be noted that the iGluR LBD observations are based on studies of isolated LBDs. In future work, we plan to calculate LBD PMFs in the context of intact receptor tetramers containing the amino-terminal and transmembrane domains. Although computationally expensive to perform, we anticipate that results of these experiments will likely add to the complexity reported here.

3.4 Materials and Methods

3.4.1 Expression in *Xenopus* Oocytes and Functional Analysis.

Full-length cDNAs for ML032222a, with R703K and E423S mutations created by overlap PCR, were sequenced and cloned into pGEMHE [35], linearized with PmeI, and used for cRNA expression using T7 polymerase (Ambion mMessage mMachine transcription kit). Defolliculated stage 5-6 *Xenopus* oocytes obtained from Ecocyte Bioscience were injected with between 0.1 and 1 ng cRNA and incubated at 18 °C in ND96 (96 mM NaCl, 2 mM KCl, 1.8 mM $CaCl_2$, 1 mM $MgCl_2$, 5 mM Hepes, 2.5 mM sodium pyruvate, gentamycin at 50 μg/mL, pH 7.6). Two electrode voltage-clamp recordings at a holding potential of -60 mV, with 3 M KCl agarose tipped electrodes of resistance 0.1-0.8 MΩ, were performed 3-5 d after injection of cRNAs. The bath

solution contained 100 mM NaCl, 1 mM KCl, and 5 mM Hepes, pH 7.5, to which $CaCl_2$ and $MgCl_2$ were added as required. Amino acids were dissolved in recording solution and applied by computer-operated solenoid valves essentially as reported previously [36].

3.4.2 Protein Expression and X-Ray Crystallography.

Synthetic genes with codon optimization for expression in Escherichia coli for the ML032222a LBD S1S2 construct with point mutations were created by overlap PCR, expressed as soluble proteins, purified to homogeneity using metal affinity and ion exchange chromatography, and crystallized as described previously [12, 14]. X-ray diffraction data collected at Southeast Regional Collaborative Access Team (SER-CAT) Advanced Photon Source (APS) was used to solve structures for the glycine bound complexes of the R703K and E423S mutant LBDs by Fourier synthesis, using the WT structure with mutant side chains truncated to CB atoms, and with alternate conformations, ligand atoms, and solvent removed from the starting model. Iterative cycles of model building and crystallographic refinement were carried out using COOT [37] and PHENIX [38] until the R_{free} value converged and Fo-Fc maps contained no interpretable features.

3.4.3 Proteolysis Protection and Binding Assays.

Purified ML032222a S1S2 and the ML032222a R703K S1S2 mutant were denatured in 4 M guanidinium, dialyzed to remove endogenous glycine, and then refolded using buffers made with HPLC grade water; apo protein for the E423S mutant was successfully prepared by dialysis without unfolding. Digestions were performed with trypsin at a ratio of 1:20 (wt/wt) at room temperature and then run on polyacrylamide gels. ITC experiments with a VP-ITC calorimeter (MicroCal) were performed at 20 °C as described previously [14].

3.4.4 Free Energy Landscapes.

The protein conformational free energy landscapes $W(\xi_1, \xi_2)$, or PMFs, were computed using an umbrella sampling simulation strategy. A 2D order parameter, (ξ_1, ξ_2), is used to describe large-scale conformational transitions in the LBD. ξ_1 and ξ_2 each specify the distance between the center of mass (COM) of an atom selection in lobe 1 and the COM of an atom selection in lobe 2. ξ_1 is the distance between the atoms N, CA, CB, C, and O of residues 498-500 (residues 92-94 of PDB ID codes 4YKI, 5CMB, and 5CMC) in lobe 1 and residues 704-705 (residues 145-146 of PDB ID codes 4YKI, 5CMB, and 5CMC) in lobe 2; ξ_2 is the distance between the same atoms of residues 423-425 (residues 17-19 of PDB ID codes 4YKI, 5CMB, and 5CMC) in lobe 1 and residues 730-731 (residues 171-172 of PDB ID codes 4YKI, 5CMB, and 5CMC) in lobe 2.

The umbrella sampling windows consist of LBD conformations positioned in 1-Å increments along ξ_1 and ξ_2. These coordinates were obtained via targeted MD simulations using CHARMM [39] initiated from the following crystal structures: 4YKI (glycine-bound WT LBD), 5CMB (glycine-bound R703K LBD), and 5CMC (glycine-bound E423S LBD). Chain B was used; residue isomer A was used from the entries where multiple options exist. Missing residues were built using the MODLOOP server [40]. During the targeted MD simulations, RMSD restraints were applied separately to each lobe to minimize intralobe structural distortions, and the glycine ligand was restrained to remain docked to Arg131 in lobe 1.

One hundred ten umbrella sampling windows were used to compute each of the PMFs. All simulations were performed using CHARMM and NAMD [41] with explicit solvent at 300 K. The all-atom potential-energy function PARAM27 for proteins [42, 43] and the TIP3P potential energy function for water [44] were used. The total simulation time for each of the apo WT, glycine-bound WT, apo R703K, glycine-bound R703K, and apo E423S PMFs is ∼400 ns. The total simulation time for the glycine-bound E423S PMF is ∼260 ns. A time step of 2 fs was used in the simulations. The number of atoms in each simulation system is ∼31,000. Crystallographic waters in each ligand-binding cleft were included in our models. Na^+ and Cl^- ions were

added in the bulk solution to give 150 mM NaCl and an electrically neutral system. Periodic boundary conditions were used with an orthorhombic cell with approximate dimensions $88 \times 64 \times 56$ Å. Harmonic biasing potentials with a force constant of 2 kcal/mol/Å^2 centered on (ξ_1, ξ_2) were used. Each PMF, $W(\xi_1, \xi_2)$, was computed using the weighted histogram analysis method (WHAM) [45, 46] to unbias and recombine the sampled distribution functions in (ξ_1, ξ_2) from all windows. A timestep of 20 fs was used in the WHAM calculations. To safeguard against ligand dissociation in open LBD conformations, an asymmetric harmonic potential (force constant of 10 kcal/mol/Å^2) was applied to the distance between the α-carboxylate oxygen atoms of the ligand and the guanidinium nitrogen atoms of Arg131. The asymmetric potential resulted in the restraint being active only when the distance exceeded 3.2 Å. This restraint is not expected to adversely affect conformational sampling because dissociation events were never observed in test simulations involving open LBD conformations.

3.4.5 Block Averaging.

The statistical uncertainty in each PMF was evaluated using the approach of block averaging [47]. For each PMF, the time series in (ξ_1, ξ_2) was divided into 10 blocks, WHAM was used to calculate a PMF from the data in each block, and then the SD in the 10 PMFs was calculated. Using 5-15 blocks all gave qualitatively similar results.

3.5 Acknowledgments

We thank Carla Glasser for technical assistance and Drs. Joy Zhao and Peter Schuck for help with analysis of ITC experiments. Data were collected at SER-CAT 22-ID beamline at the APS, Argonne National Laboratory. Use of the APS was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38. This study used the high-performance computational capabilities of the Biowulf Linux cluster at the NIH and resources provided by the Maryland Advanced Research Computing Center at Johns Hopkins University. This work was supported by the intramural research program of The Eunice

Kennedy Shriver National Institute of Child Health and Human Development, NIH, Department of Health and Human Services (to Mark L. Mayer) and NIH Grant R01GM094495 (to Albert Y.Lau).

Fig. 3-1: Salt bridge mutants speed recovery from desensitization (A) Crystal structure of the WT ML032222a iGluR LBD (PDB ID code 4YKI). Lobe 1 is colored gold and lobe 2 colored pale green; labels indicate secondary structure helix assignments. (B) Zoomed view of the entrance to the WT ML032222a binding site showing the interdomain salt bridge formed by the side chains of Glu423 in lobe 1 (gold) and Arg703 in lobe 2 (pale green). (C) WT ML032222a responses to paired 5-s applications of glycine in which the interval between the first and second application was varied from 10 s to 5 min to measure the rate of recovery from desensitization. (D) ML032222a R703K mutant responses to paired 5-s applications of glycine in which the interval between the first and second pulse was varied from 10 to 90 s. (E) ML032222a E423S mutant responses to paired 5-s applications of glycine in which the interval between the first and second pulse was varied from 2 to 5 s. (F) Bar plot showing the rate constant for recovery from desensitization for WT ML032222a and the R703K and E423S mutants; error bars show SEM for three to four observations.

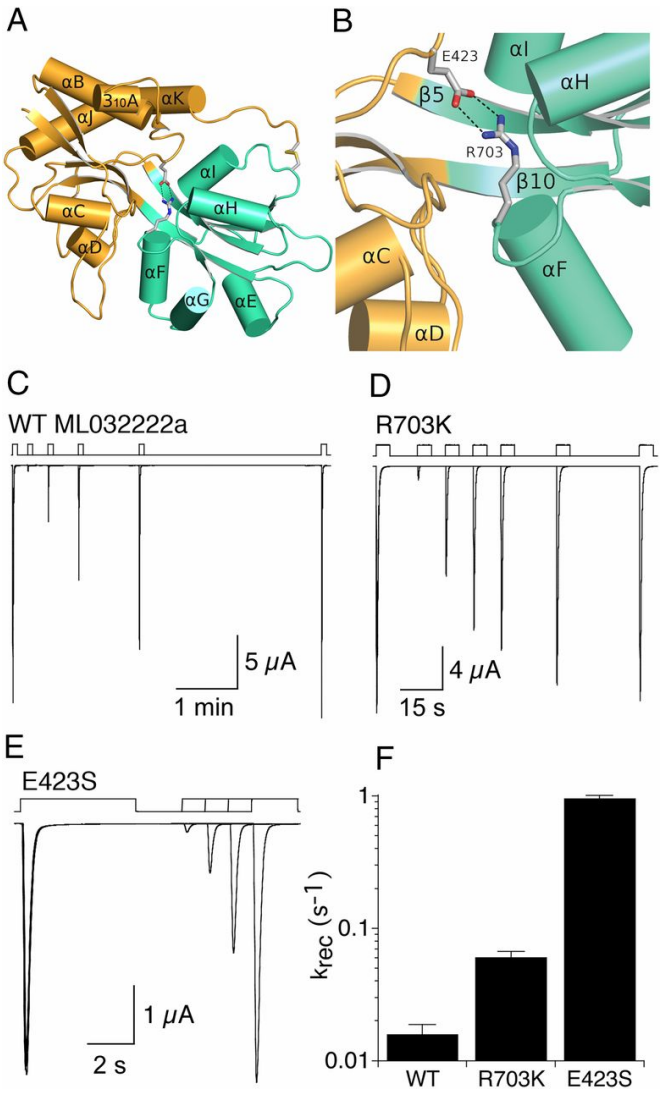
A
αB
310A
αK
αJ
αI
αH
αC
αD
αF
αG
αE
B
E423
αI
αH
β5
R703
β10
αC
αF
αD
C
WT ML032222a
5 μA
1 min
D
R703K
4 μA
15 s
E
E423S
1 μA
2 s
F
1
0.1
0.01
krec (s-1)
WT
R703K
E423S

Fig. 3-1

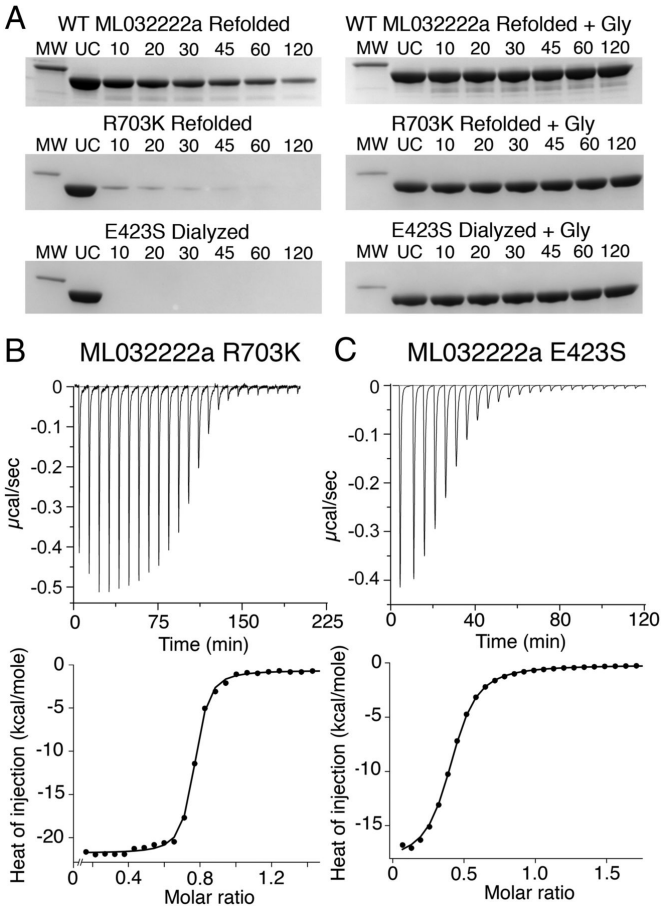

Fig. 3-2: **Salt bridge mutants lower affinity for glycine.** (A) Proteolysis protection assays for refolded WT ML032222a S1S2 (Top); refolded R703K mutant (Middle); and exhaustively dialyzed E423S mutant (Bottom). (Left) Coomassie blue stained SDS/PAGE experiments show the time course of digestion by trypsin; (Right) protection by 1 mM glycine; lanes show a 31-kDa marker (MW), uncut protein (UC), and samples at the indicated times in minutes after addition of trypsin. (B) Titration of refolded ML032222a R703K by glycine analyzed by ITC, with raw (Top) and integrated (Bottom) data fit with a binding isotherm of $K_d = 28$ nM. (C) Titration of ML032222a E423S by glycine analyzed by ITC, with raw (Top) and integrated (Bottom) data fit with a binding isotherm of $K_d = 2.5$ μM.

Fig. 3-3: Salt bridge mutant crystal structures. (A) Crystal structure stereoview of the WT ML032222a binding pocket, with a 1.21-Å resolution 2mFo-DFc electron density map contoured at 2σ for ligand and protein, and 3σ for water molecules; secondary structure elements for the S1 and S2 segments are colored gold and green, respectively; side chains involved in ligand binding and interdomain contacts are drawn in stick representation, with ion pair and hydrogen bond contacts drawn as dashed lines; water molecules flanking the entrance to the binding site are shown as red spheres; note the hydrogen bond contact of W9 with the Glu423 side chain; the surface of the ligand binding pocket is shaded in gray. (B) Crystal structure stereoview of the ML032222a R703K mutant binding pocket colored as above with a 1.34-Å resolution 2mFo-DFc electron density map contoured at 2σ; note the change in side chain torsion angle for Glu423, movement of water molecules W7 and W8, and interposition of W9 between the side chains of Glu747 and Asp497. (C) Crystal structure stereoview of the ML032222a E423S mutant binding pocket colored as above with a 1.28-Å resolution 2mFo-DFc electron density map contoured at 2σ for ligand and protein and 2.5σ for water molecules; note movement of W7 into the position occupied by the E423 γ-COOH group in the WT receptor and the accompanying movement of W8.

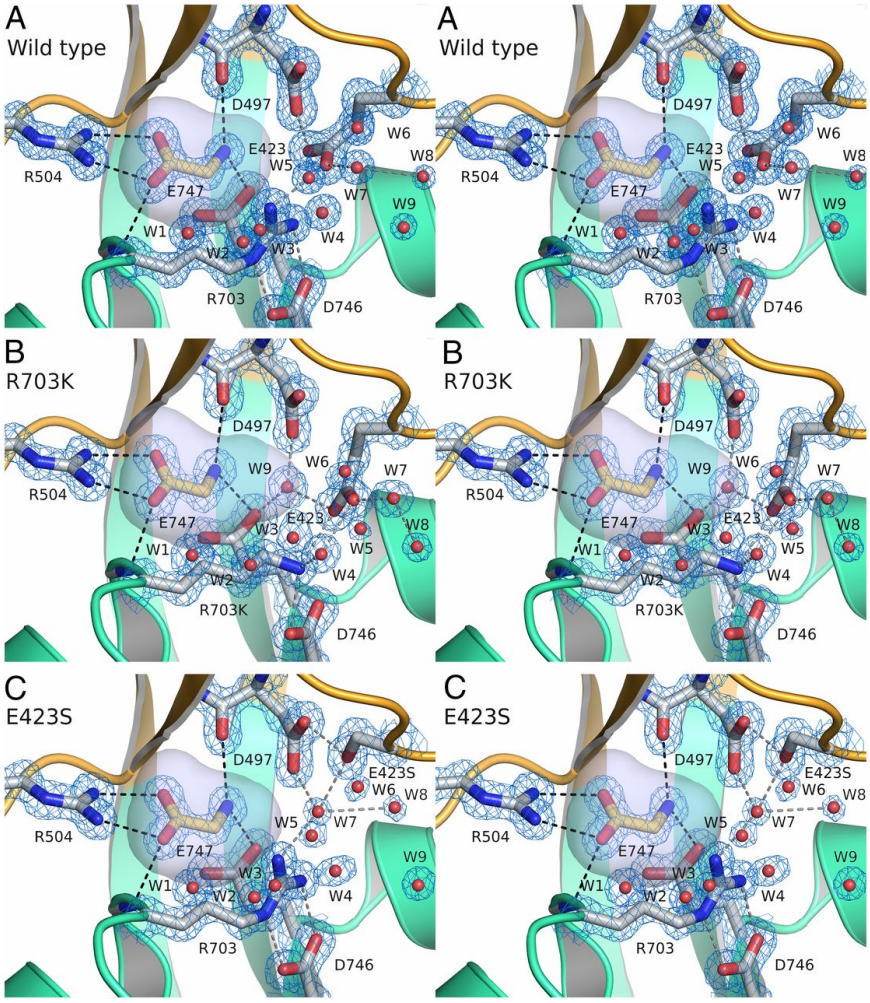
A
Wild type
D497
E423
W5
W6
W8
R504
E747
W7
W9
W1
W4
W2
W3
R703
D746
A
Wild type
D497
E423
W5
W6
W8
R504
E747
W7
W9
W1
W4
W2
W3
R703
D746
B
R703K
D497
W9
W6
W7
R504
E747
E423
W8
W3
W1
W5
W4
W2
R703K
D746
B
R703K
D497
W9
W6
W7
R504
E747
E423
W8
W3
W1
W5
W4
W2
R703K
D746
C
E423S
D497
E423S
W6
W8
W5
W7
R504
E747
W9
W3
W1
W2
W4
R703
D746
C
E423S
D497
E423S
W6
W8
W5
W7
R504
E747
W9
W3
W1
W2
W4
R703
D746

Fig. 3-3

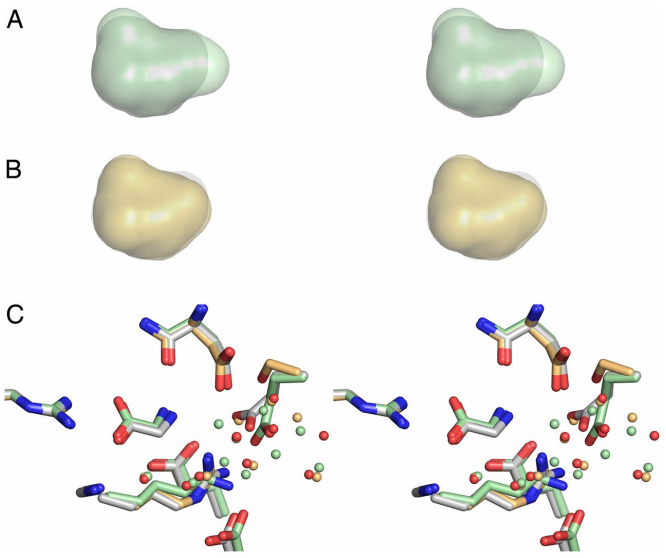

Fig. 3-4: **Comparison of LBD cavity sizes and water molecule networks in the WT, R703K, and E423S mutants.** (A) Stereoview of a superposition of the LBD cavity for WT (gray) and R703K (pale green). (B) Stereoview of a superposition of the LBD cavity for WT (gray) and E423S (pale orange). (C) Stereoview of superpositions using domain 1 main chain coordinates for residues forming the glycine binding site, with carbon atoms and water molecules colored pale green for R703K and pale orange for E423S, respectively; for WT, carbon atoms are gray with water molecules colored red.

Fig. 3-5: Conformational free energy landscapes.(A) The order parameter (ξ_1, ξ_2) used to describe large-scale conformational transitions in the ML032222a LBD. ξ_1 and ξ_2 each represent the distance (dashed line) between the upper and lower lobes of the protein at the centers of mass of the highlighted atoms. Disulfide bonds are shown in red. (B) Free energy landscapes for the WT apo state (Top) and glycine complex (Bottom) plotted as 2D PMFs (Left) and 1D PMFs (Right). Contour lines correspond to a difference of 1 kcal/mol, with cooler colors being lower in free energy (see the color bar). **Fig. 3-7** shows the SDs for the 2D PMFs, as determined by block averaging. For the 1D PMFs, the red-shaded region indicates the SD of the free energy. The locations of crystal structures are indicated by yellow dots. (C) Free energy landscapes for the R703K LBD. (D) Free energy landscapes for the E423S LBD.

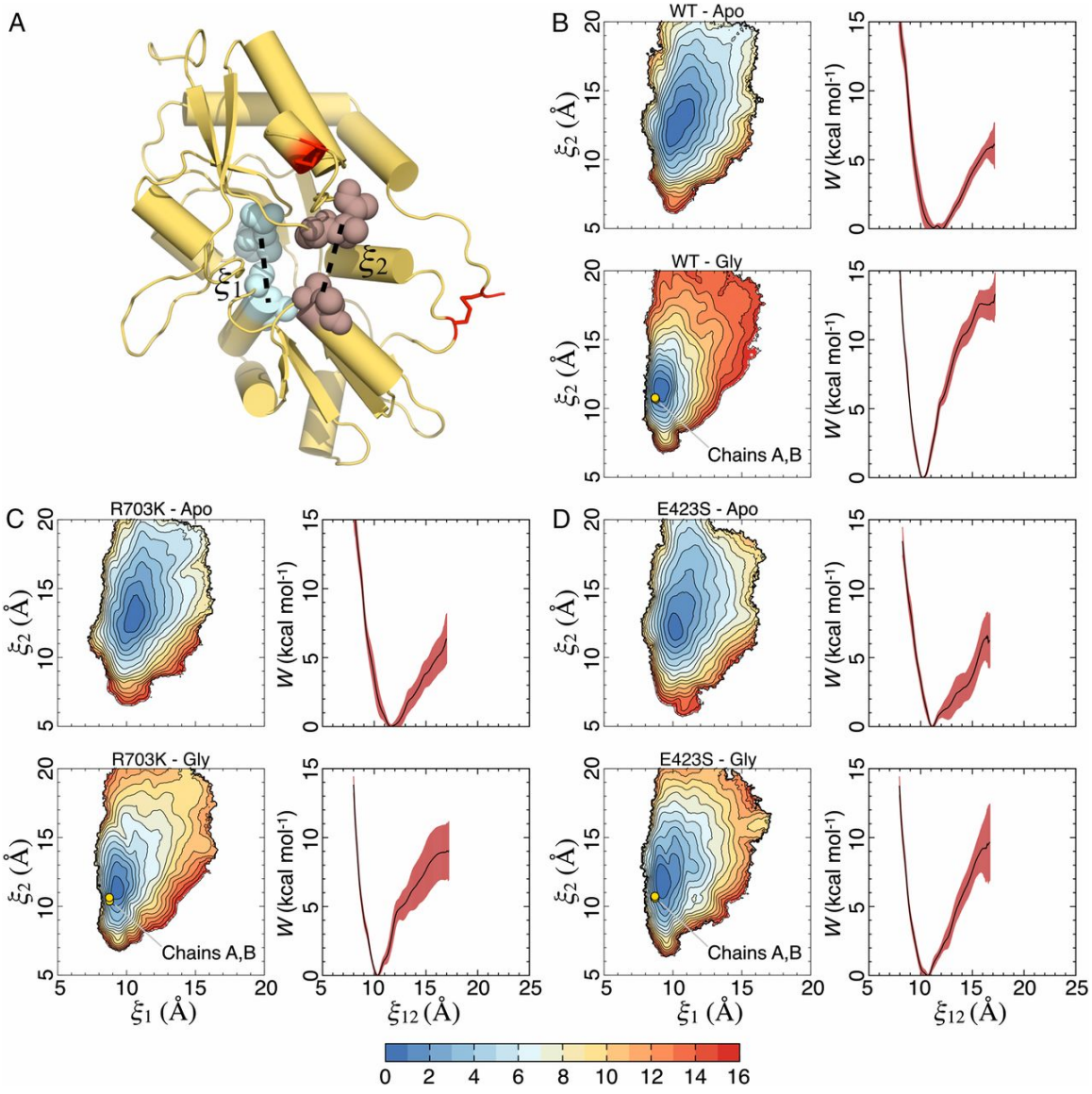
A
B
WT - Apo
WT - Gly
Chains A,B
C
R703K - Apo
R703K - Gly
Chains A,B
D
E423S - Apo
E423S - Gly
Chains A,B
ξ1
ξ2
ξ2 (Å)
W (kcal mol-1)
ξ1 (Å)
ξ12 (Å)
0 2 4 6 8 10 12 14 16

Fig. 3-5

Fig. 3-6: Comparison of free energy basins for the apo and glycine complex LBDs. (A) The lowest three contour lines from the 2D PMFs in **Fig. 3-5** are shown for the WT LBD. Apo is shown in black, and the glycine complex is shown in red. (B) Contours for the R703K LBD. (C) Contours for the E423S LBD.

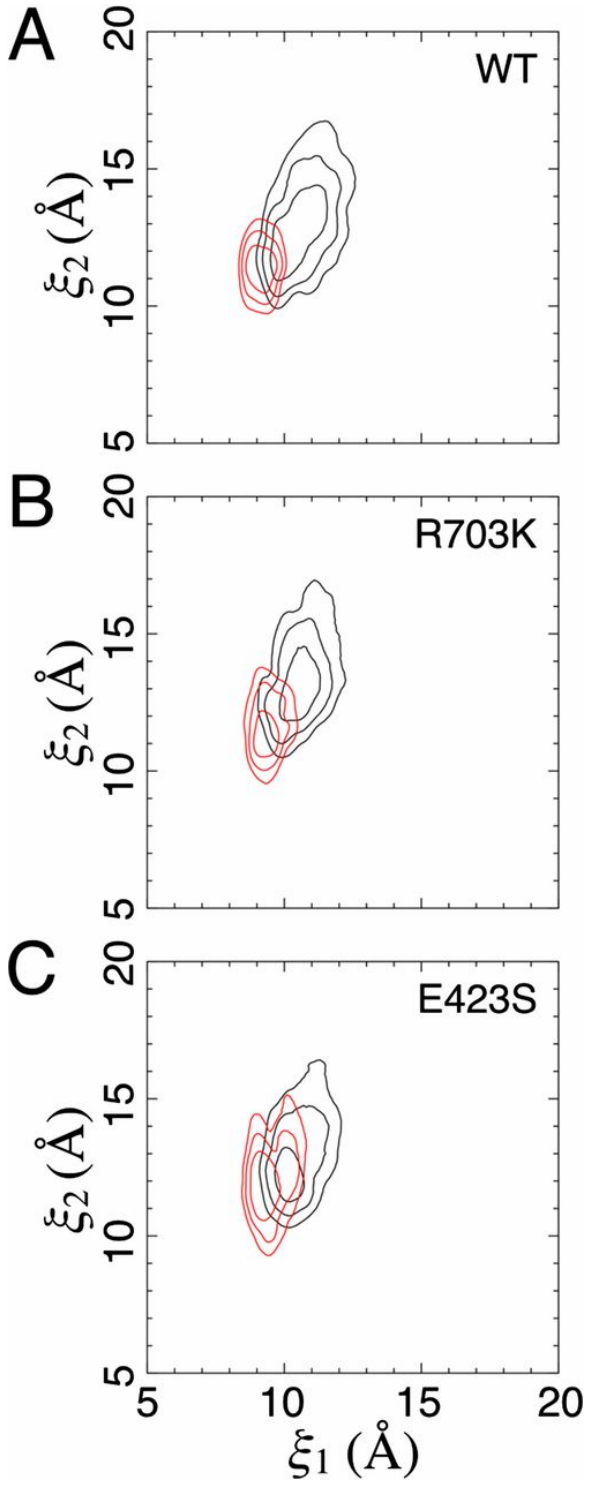

Fig. 3-6

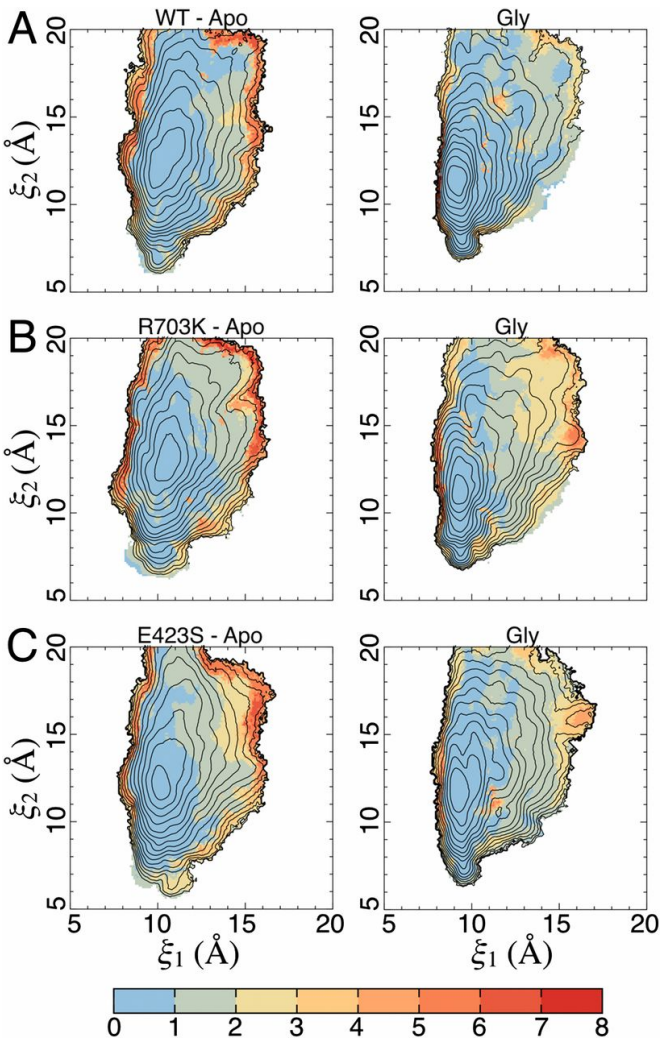

Fig. 3-7: **The SD of the 2D PMFs .** (A) SD in **Fig. 3-5**B. (Left) The apo state. (Right) The glycine complex. SDs are shown in kcal/mol (see the color bar), as determined by block averaging. The contour lines are taken from the 2D PMFs (**Fig. 3-5**). (B) SD in **Fig. 3-5**C. (C) SD in **Fig. 3-5**D.

Fig. 3-8: Salt bridge conformational dynamics. (A) Interactions between Arg703 and Glu423 for the WT LBD as a function of cleft closure. (Left) The apo state. (Right) The glycine complex. Average distances are shown in Ångstroms (see the color bar); SDs are shown in **Fig. 3-9** A and B. The contour lines are taken from the 2D PMFs (**Fig. 3-5**). (B) Interactions between Lys703 and Glu423 for the R703K mutant LBD. (C) Distance between the Glu423 Cδ atom and the Asp497 Cγ atom for the WT LBD as a function of cleft closure. (Left) The apo state. (Right) The glycine complex. Average distances are shown in Ångstroms; SDs are shown in **Fig. 3-9** C and D. The contour lines are taken from the 2D PMFs (**Fig. 3-5**). (D) Distance between Glu423 and Asp497 for the R703K mutant LBD.

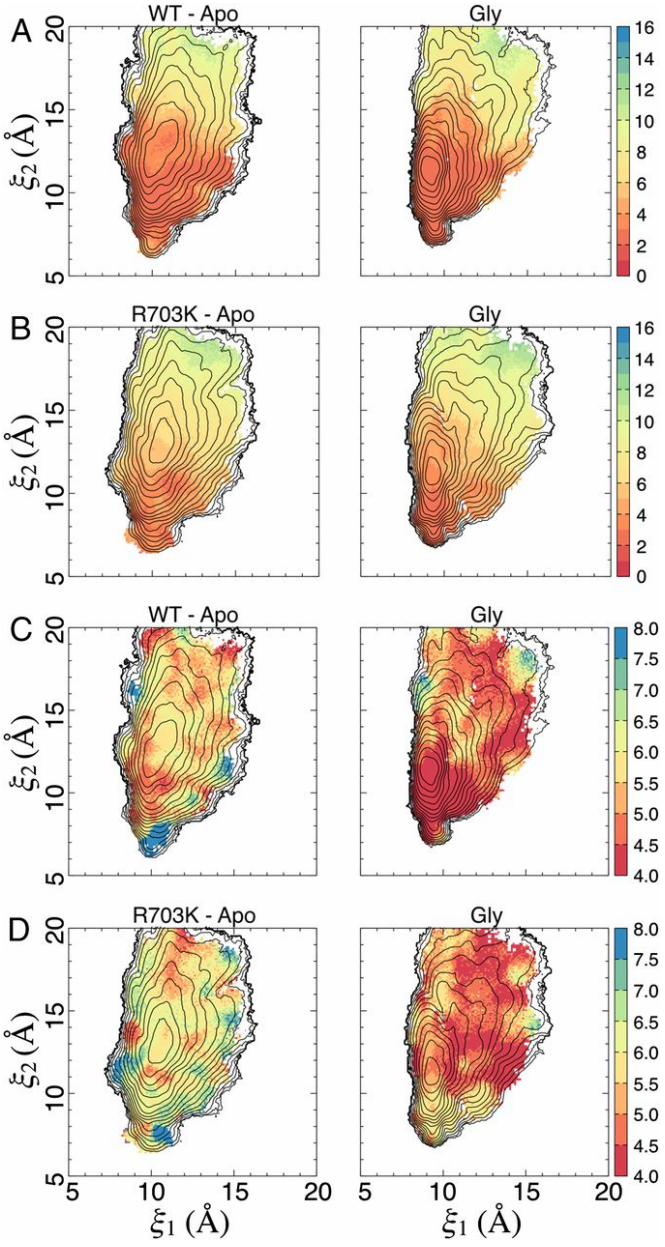
A
WT - Apo
Gly
B
R703K - Apo
Gly
C
WT - Apo
Gly
D
R703K - Apo
Gly
ξ2 (Å)
ξ1 (Å)
5
10
15
20
16
14
12
10
8
6
4
2
0
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0

Fig. 3-8

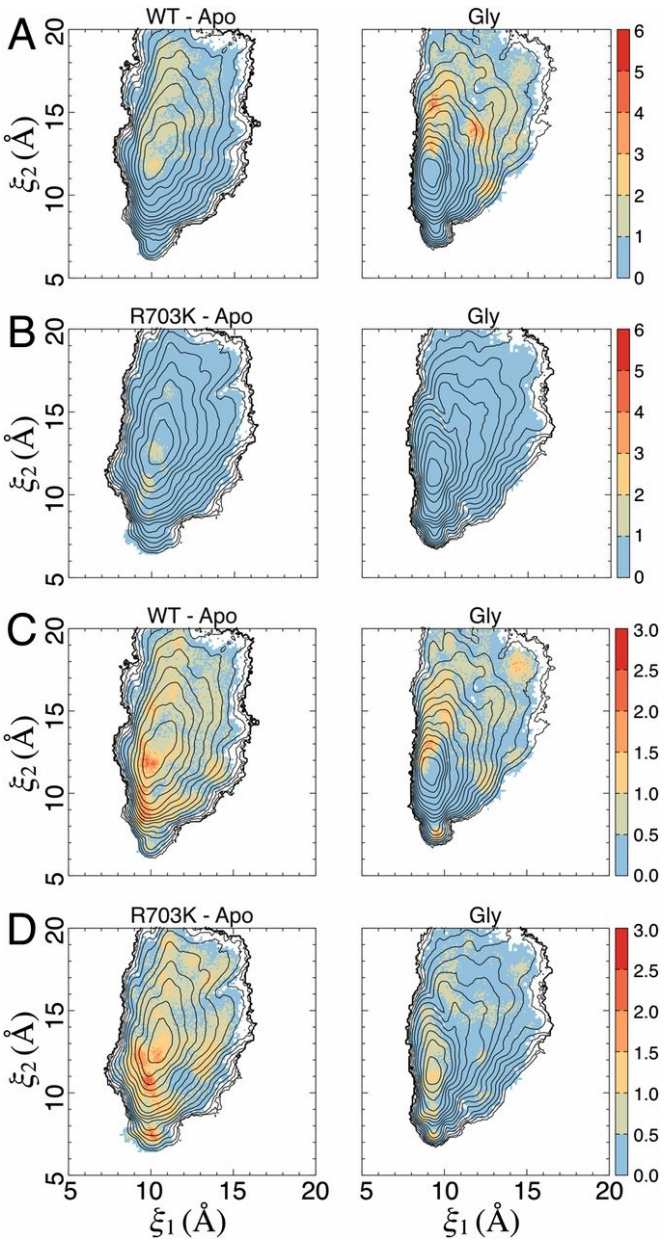

Fig. 3-9: **The SD of distance measurements involving the salt bridge residues (Fig. 3-8)**. (A) SD in **Fig. 3-8**A. (Left) The apo state. (Right) The glycine complex. SDs are shown in Ångstroms (see the color bars). The contour lines are taken from the 2D PMFs (**Fig. 3-5**). (B) SD in **Fig. 3-8**B. (C) SD in **Fig. 3-8**C. (D) SD in **Fig. 3-8**D.

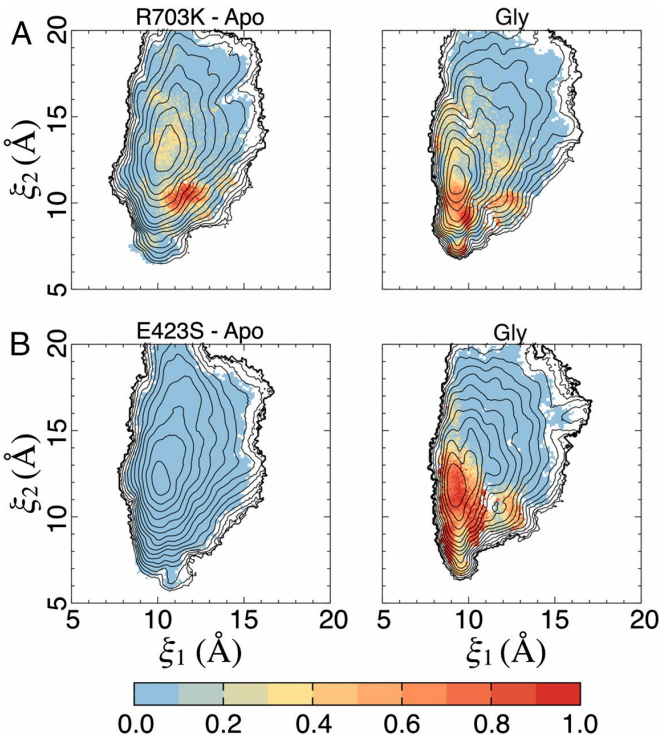

Fig. 3-10: **Occupancy of binding pocket water molecules as a function of cleft closure.** (A) For the R703K mutant LBD, the binding pocket water corresponds to W9 in **Fig. 3-3**B and is coordinated by Glu423, Asp497, and Glu747. (Left) The apo state. (Right) The glycine complex. Cooler colors indicate lower occupancy. The contour lines are taken from the 2D PMFs (**Fig. 3-5**). (B) For the E423S mutant LBD, the water corresponds to W7 in **Fig. 3-3**C and is coordinated by Ser423, Asp497, and Arg703.

Fig. 3-11: Glycine accessibility as a function of cleft closure. (A) Tunnels that allow a glycine ligand to access its binding pocket from bulk solvent, found using CAVER Analyst, are shown in green. The primary access tunnel to the pocket is in the center; the lateral tunnels are crevices that feed into the central tunnel. Bottleneck residues for the central tunnel are Phe469 and Arg703, shown in red. Glu423 also lines the central tunnel and is shown in magenta. A minimum distance of ∼10 Åbetween the side chains of residues 469 and 703, measured at Cγ atoms, is required for glycine passage. Disulfide bonds are shown in orange. The example LBD conformation shown corresponds to $(\xi_1, \xi_2) \sim$ (13, 16 Å). In this conformation, the Arg703-Glu423 salt bridge is broken. (B) Distance between the side chains of Phe469 and Arg703, measured as described in A, for the WT BD. (Left) The apo state. (Right) The glycine complex. Average distances are shown in units of Ångstroms (see the color bar); SDs are shown in **Fig. 3-12**. The dashed box indicates the region of integration for open LBD conformations. The contour lines are taken from the 2D PMFs (**Fig. 3-5**). (C) Distance between the side chains of Phe469 and Lys703 for the R703K mutant LBD. (D) Distance between the side chains of Phe469 and Arg703 for the E423S mutant LBD.

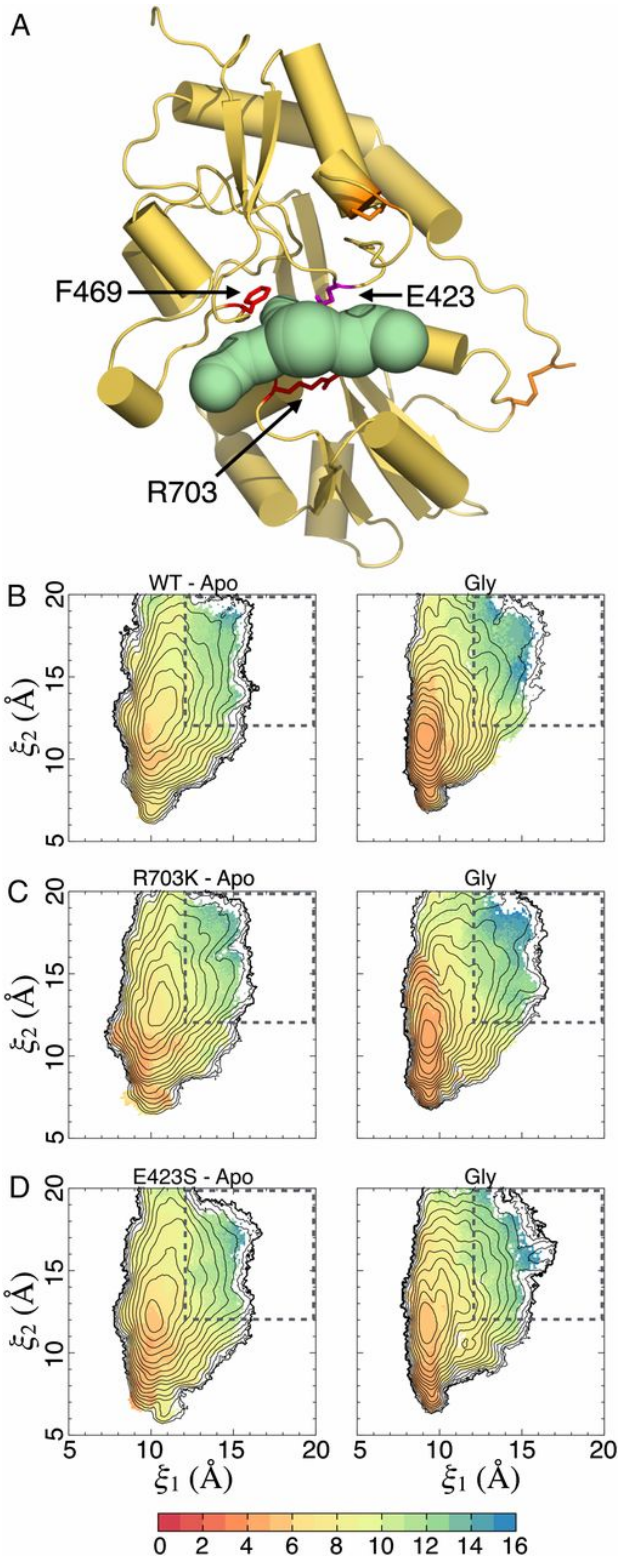
A
F469
E423
R703
B
WT - Apo
Gly
C
R703K - Apo
Gly
D
E423S - Apo
Gly
ξ2 (Å)
ξ1 (Å)
5
10
15
20
0
2
4
6
8
10
12
14
16

Fig. 3-11

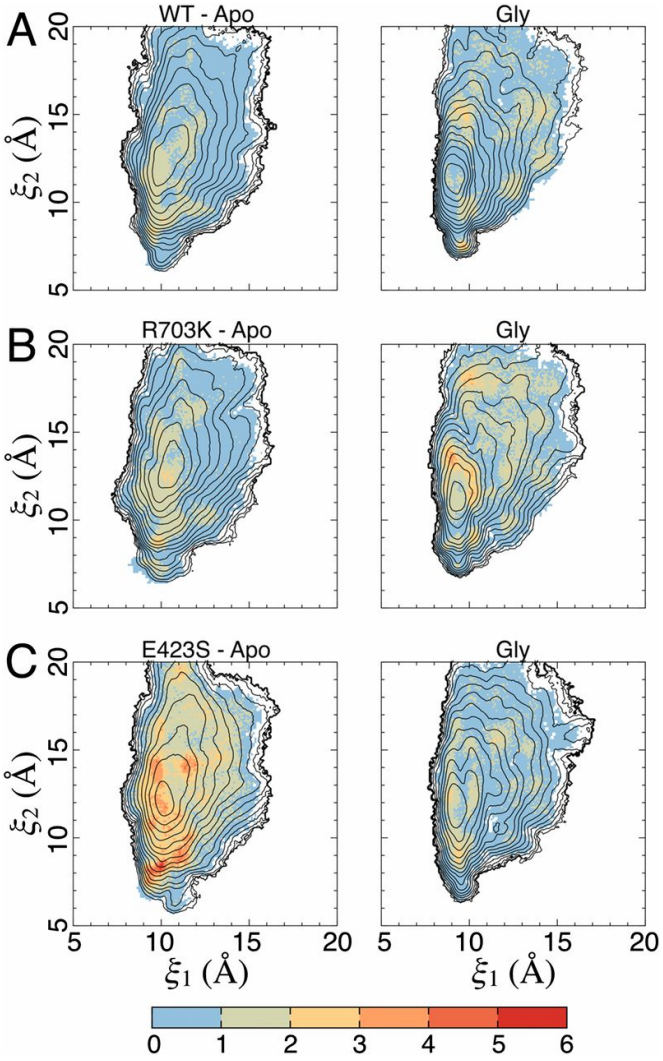

Fig. 3-12: **The SD of distance measurements between Phe469 and Arg/Lys703 (Fig. 3-11).** (A) SD in **Fig. 3-11**B. (Left) The apo state. (Right) The glycine complex. SDs are shown in Ångstroms (see the color bar). The contour lines are taken from the 2D PMFs (**Fig. 3-5**). (B) SD in **Fig. 3-11**C. (C) SD in **Fig. 3-11**D.

	R703K	E423S
Data collection		
Space group	$P2_1$	$P2_1$
Unit cell *a, b, c* (Å)	45.7, 123.4, 54.3	45.9, 123.1, 54.5
α, β, γ	90, 112.2, 90	90, 111.7, 90
Number per a.u.	2	2
Wavelength (Å)	1.0000	1.0000
Resolution (Å)*	40.0–1.34	40–1.28 (1.30)
Unique observations	124,471	144,111
Mean redundancy†	4.4 (2.6)	4.5 (3.3)
Completeness (%)†	97.8 (72.9)	96.6 (81.9)
R_{merge}†,‡,	0.045 (0.365)	0.040 (0.323)
R_{pim}†,§	0.024 (0.261)	0.021 (0.201)
I/σ(I)†	31.1 (2.4)	34.9 (3.5)
Refinement		
Resolution (Å)	39.0–1.34	39.1–1.28
Protein atoms (AC)¶	4,109 (107)	4,217 (217)
Ligand atoms	10	10
Mg/SO_4 ions	1/1	1/1
Water atoms	780	664
R_{work}/R_{free} (%)‖	14.3/16.7	13.4/15.7
RMSD		
Bond lengths (Å)	0.009	0.027
Bond angles °	1.24	1.27
Mean B-values (Å^2)		
Protein overall	17.8	20.4
MC/SC#	15.9/19.7	19.6/22.1
Ligand	9.3	11.6
Mg/SO_4 ions	8.9/47.4	17.4/55.6
Water	31.0	30.7
Ramachandran %**	97.7/0.60	98.8/1.3

*Values in parentheses indicate the low-resolution limit for the highest-resolution shell of data.

†Values in parentheses indicate statistics for the highest-resolution shell of data.

‡$R_{merge} = (\sum_{hkl}\sum_j |I_{hkl,j} - \langle I_{hkl}\rangle|)/(\sum_{hkl}\sum_j I_{hkl,j})$, where $\langle I_{hkl}\rangle$ is the mean I_{hkl} over symmetry–equivalent reflections.

§$R_{pim} = (\sum_{hkl}\sqrt{1/(n-1)}\sum_{j=1}^{n} |I_{hkl,j} - \langle I_{hkl}\rangle|)/(\sum_{hkl}\sum_j I_{hkl,j})$, where $\langle I_{hkl}\rangle$ is the mean I_{hkl} over symmetry-equivalent reflections.

¶Alternate conformations.

‖$R_{work} = (\sum_{hkl}|F_o - F_c|)/(\sum_{hkl}F_o)$, where F_o and F_c denote observed and calculated structure factors, respectively; 5% of the reflections were set aside for the calculation of the R_{free} value.

#Main chain/side chain.

**Preferred/disallowed conformations.

Table 3.1: Data collection and refinement statistics

	WT (kcal/mol)	R703K (kcal/mol)	E423S (kcal/mol)
$\Delta G_{\text{conf,Gly}}$	9.474	5.966	4.977
$-\Delta G_{\text{conf,Apo}}$	−2.543	−2.862	−2.848
$\Delta\Delta G_{\text{conf}}$	6.931	3.104	2.129

Table 3.2: **Free energy of conformational change in the LBD**

References

[1] Traynelis, S. F., Wollmuth, L. P., McBain, C. J., Menniti, F. S., Vance, K. M., Ogden, K. K., Hansen, K. B., Yuan, H., Myers, S. J., and Dingledine, R. Glutamate Receptor Ion Channels: Structure, Regulation, and Function. *Pharmacological Reviews*, 62(3):405–496, September 2010.

[2] Karakas, E. and Furukawa, H. Crystal structure of a heterotetrameric NMDA receptor ion channel. *Science*, 344(6187):992–997, May 2014.

[3] Lee, C.-H., Lu, W., Michel, J. C., Goehring, A., Du, J., Song, X., and Gouaux, E. NMDA receptor structures reveal subunit arrangement and pore architecture. *Nature*, 511(7508):191–197, July 2014.

[4] Yao, Y., Harrison, C. B., Freddolino, P. L., Schulten, K., and Mayer, M. L. Molecular mechanism of ligand recognition by NR3 subtype glutamate receptors. *The EMBO Journal*, 27(15):2158–2170, August 2008.

[5] Monyer, H., Sprengel, R., Schoepfer, R., Herb, A., Higuchi, M., Lomeli, H., Burnashev, N., Sakmann, B., and Seeburg, P. H. Heteromeric NMDA Receptors: Molecular and Functional Distinction of Subtypes. *Science*, 256(5060):1217–1221, May 1992.

[6] Moriyoshi, K., Masu, M., Ishii, T., Shigemoto, R., Mizuno, N., and Nakanishi, S. Molecular cloning and characterization of the rat NMDA receptor. *Nature*, 354(6348):31–37, November 1991.

[7] Burnashev, N. and Szepetowski, P. NMDA receptor subunit mutations in neurodevelopmental disorders. *Current Opinion in Pharmacology*, 20:73–82, February 2015.

[8] Kumar, J. and Mayer, M. L. Functional insights from glutamate receptor ion channel structures. *Annual Review of Physiology*, 75:313–337, 2013.

[9] Furukawa, H., Singh, S. K., Mancusso, R., and Gouaux, E. Subunit arrangement and function in NMDA receptors. *Nature*, 438(7065):185–192, November 2005.

[10] Furukawa, H. and Gouaux, E. Mechanisms of activation, inhibition and specificity: crystal structures of the NMDA receptor NR1 ligandâĂŘbinding core. *The EMBO Journal*, 22(12):2873–2885, June 2003.

[11] Armstrong, N. and Gouaux, E. Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron*, 28(1):165–181, October 2000.

[12] Mayer, M. L. Crystal structures of the GluR5 and GluR6 ligand binding cores: molecular mechanisms underlying kainate receptor selectivity. *Neuron*, 45(4):539–552, February 2005.

[13] Mayer, M. L. Glutamate receptors at atomic resolution. *Nature*, 440(7083):456–462, March 2006.

[14] Alberstein, R., Grey, R., Zimmet, A., Simmons, D. K., and Mayer, M. L. Glycine activated ion channel subunits encoded by ctenophore glutamate receptor genes. *Proceedings of the National Academy of Sciences*, 112(44):E6048–E6057, November 2015.

[15] Robert, A., Armstrong, N., Gouaux, J. E., and Howe, J. R. AMPA Receptor Binding Cleft Mutations That Alter Affinity, Efficacy, and Recovery from Desensitization. *Journal of Neuroscience*, 25(15):3752–3762, April 2005.

[16] Weston, M. C., Gertler, C., Mayer, M. L., and Rosenmund, C. Interdomain Interactions in AMPA and Kainate Receptors Regulate Affinity for Glutamate. *Journal of Neuroscience*, 26(29):7650–7658, July 2006.

[17] Lau, A. Y. and Roux, B. The Free Energy Landscapes Governing Conformational Changes in a Glutamate Receptor Ligand-Binding Domain. *Structure*, 15(10):1203–1214, October 2007.

[18] Lau, A. Y. and Roux, B. The hidden energetics of ligand binding and activation in a glutamate receptor. *Nature Structural & Molecular Biology*, 18(3):283–287, March 2011.

[19] Yao, Y., Belcher, J., Berger, A. J., Mayer, M. L., and Lau, A. Y. Conformational analysis of NMDA receptor GluN1, GluN2, and GluN3 ligand-binding domains reveals subtype-specific characteristics. *Structure (London, England: 1993)*, 21(10):1788–1799, October 2013.

[20] Landes, C. F., Rambhadran, A., Taylor, J. N., Salatan, F., and Jayaraman, V. Structural landscape of isolated agonist-binding domains from single AMPA receptors. *Nature Chemical Biology*, 7(3):168–173, March 2011.

[21] Dolino, D. M., Cooper, D., Ramaswamy, S., Jaurich, H., Landes, C. F., and Jayaraman, V. Structural Dynamics of the Glycine-binding Domain of the N-Methyl-d-Aspartate Receptor. *Journal of Biological Chemistry*, 290(2):797–804, January 2015.

[22] Li, H., Robertson, A. D., and Jensen, J. H. Very fast empirical prediction and rationalization of protein pKa values. *Proteins*, 61(4):704–721, December 2005.

[23] Kozlikova, B., Sebestova, E., Sustr, V., Brezovsky, J., Strnad, O., Daniel, L., Bednar, D., Pavelka, A., Manak, M., Bezdeka, M., Benes, P., Kotry, M., Gora, A., Damborsky, J., and Sochor, J. CAVER Analyst 1.0: graphic tool for interactive visualization and analysis of tunnels and channels in protein structures. *Bioinformatics*, 30(18):2684–2685, September 2014.

[24] Moroz, L. L., Kocot, K. M., Citarella, M. R., Dosung, S., Norekian, T. P., Povolotskaya, I. S., Grigorenko, A. P., Dailey, C., Berezikov, E., Buckley, K. M., Ptitsyn, A., Reshetov, D., Mukherjee, K., Moroz, T. P., Bobkova, Y., Yu, F., Kapitonov, V. V., Jurka, J., Bobkov, Y. V., Swore, J. J., Girardo, D. O., Fodor, A., Gusev, F., Sanford, R., Bruders, R., Kittler, E., Mills, C. E., Rast, J. P., Derelle, R., Solovyev, V. V., Kondrashov, F. A., Swalla, B. J., Sweedler, J. V., Rogaev, E. I., Halanych, K. M., and Kohn, A. B. The ctenophore genome and the evolutionary origins of neural systems. *Nature*, 510(7503):109–114, June 2014.

[25] Ryan, J. F., Pang, K., Schnitzler, C. E., Nguyen, A.-D., Moreland, R. T., Simmons, D. K., Koch, B. J., Francis, W. R., Havlak, P., Smith, S. A., Putnam,

N. H., Haddock, S. H. D., Dunn, C. W., Wolfsberg, T. G., Mullikin, J. C., Martindale, M. Q., and Baxevanis, A. D. The Genome of the Ctenophore Mnemiopsis leidyi and Its Implications for Cell Type Evolution. *Science*, 342(6164):1242592, December 2013.

[26] Kalbaugh, T. L., VanDongen, H. M. A., and VanDongen, A. M. J. Ligand-Binding Residues Integrate Affinity and Efficacy in the NMDA Receptor. *Molecular Pharmacology*, 66(2):209–219, August 2004.

[27] Maier, W., Schemm, R., Grewer, C., and Laube, B. Disruption of interdomain interactions in the glutamate binding pocket affects differentially agonist affinity and efficacy of N-methyl-D-aspartate receptor activation. *The Journal of Biological Chemistry*, 282(3):1863–1872, January 2007.

[28] Zhang, W., Cho, Y., Lolis, E., and Howe, J. R. Structural and Single-Channel Results Indicate That the Rates of Ligand Binding Domain Closing and Opening Directly Impact AMPA Receptor Gating. *Journal of Neuroscience*, 28(4):932–943, January 2008.

[29] MacLean, D. M., Ramaswamy, S. S., Du, M., Howe, J. R., and Jayaraman, V. Stargazin promotes closure of the AMPA receptor ligand-binding domain. *The Journal of General Physiology*, 144(6):503–512, December 2014.

[30] Le Meur, K., Galante, M., Angulo, M. C., and Audinat, E. Tonic activation of NMDA receptors by ambient glutamate of non-synaptic origin in the rat hippocampus. *The Journal of Physiology*, 580(Pt. 2):373–383, April 2007.

[31] Papouin, T., Ladepeche, L., Ruel, J., Sacchi, S., Labasque, M., Hanini, M., Groc, L., Pollegioni, L., Mothet, J.-P., and Oliet, S. H. R. Synaptic and extrasynaptic NMDA receptors are gated by different endogenous coagonists. *Cell*, 150(3):633–646, August 2012.

[32] Bermejo, G. A., Strub, M.-P., Ho, C., and Tjandra, N. Ligand-free open-closed transitions of periplasmic binding proteins: the case of glutamine-binding protein. *Biochemistry*, 49(9):1893–1902, March 2010.

[33] Tang, C., Schwieters, C. D., and Clore, G. M. Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature*, 449(7165):1078–1082, October 2007.

[34] Bucher, D., Grant, B. J., Markwick, P. R., and McCammon, J. A. Accessing a hidden conformation of the maltose binding protein using accelerated molecular dynamics. *PLoS computational biology*, 7(4):e1002034, April 2011.

[35] Liman, E. R., Tytgat, J., and Hess, P. Subunit stoichiometry of a mammalian K+ channel determined by construction of multimeric cDNAs. *Neuron*, 9(5):861–871, November 1992.

[36] Panchenko, V. A., Glasser, C. R., and Mayer, M. L. Structural Similarities between Glutamate Receptor Channels and K+ Channels Examined by Scanning Mutagenesis. *The Journal of General Physiology*, 117(4):345–360, April 2001.

[37] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. Features and development of Coot. *Acta Crystallographica. Section D, Biological Crystallography*, 66(Pt 4):486–501, April 2010.

[38] Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica. Section D, Biological Crystallography*, 66(Pt 2):213–221, February 2010.

[39] Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera,

K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. CHARMM: the biomolecular simulation program. *Journal of Computational Chemistry*, 30(10):1545–1614, July 2009.

[40] Fiser, A. and Sali, A. ModLoop: automated modeling of loops in protein structures. *Bioinformatics*, 19(18):2500–2501, December 2003.

[41] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802, December 2005.

[42] MacKerell, A. D., Feig, M., and Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry*, 25(11):1400–1415, August 2004.

[43] MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry. B*, 102(18):3586–3616, April 1998.

[44] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, July 1983.

[45] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on

biomolecules. I. The method. *Journal of Computational Chemistry*, 13(8):1011–1021, October 1992.

[46] Souaille, M. and Roux, B. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer Physics Communications*, 135:40–57, March 2001.

[47] Zhu, F. and Hummer, G. Convergence and error estimation in free energy calculations using the weighted histogram analysis method. *Journal of Computational Chemistry*, 33(4):453–465, February 2012.

Chapter 4 - Neurotransmitter funneling optimizes glutamate receptor kinetics

Reproduced in part from a manuscript submitted to:

Neuron

by

Alvin Yu, Hector Salazar, Andrew J.R. Plested, and Albert Y. Lau

Ionotropic glutamate receptors (iGluRs) are found throughout the brain and are essential for excitatory synaptic transmission. The receptor is activated when the neurotransmitter glutamate binds to bilobed ligand-binding domains (LBDs), but very little is known about how glutamate reaches its recessed binding pocket and how ligand binding is coupled to protein conformational changes. Here we report the process of glutamate binding to a prototypical glutamate receptor, GluA2, in atomistic detail using unbiased molecular simulations. Charged residues on the surface of the LBD form binding pathways that facilitate glutamate binding by effectively reducing a three- dimensional diffusion process to a spatially-constrained two-dimensional one. Free energy calculations identify surface residues that metastably bind glutamate and help funnel it into the binding pocket. Surprisingly, these simulations also reveal that glutamate can bind in an inverted conformation and also reorient while in the binding pocket. Electrophysiological recordings demonstrate that eliminating these transient binding sites reveals a selective and unique kinetic signature of slowed activation and deactivation, consistent with slower glutamate binding and unbinding. These results suggest that binding pathways have evolved to optimize rapid responses of AMPA-type glutamate receptors at central nervous system synapses.

4.1 Background

The "clock speed" of the brain is set at $\sim$1 kHz by the width of action potentials, the release rate of synaptic vesicles, and the duration of synaptic potentials [1, 2]. Fast activation and deactivation of synaptic neurotransmitter receptors is therefore essential for normal signalling in the nervous system. One of the fastest-operating receptors is the AMPA-type ionotropic glutamate receptor (iGluR) [3], a ligand-gated ion channel found throughout the brain [4]. Activation of these receptors is the final step in a cascade that allows cells releasing glutamate to excite downstream target neurons with millisecond precision [5]. Each receptor has four binding sites for glutamate that resemble clamshells, termed ligand-binding domains (LBDs). The four LBDs in each receptor assemble as a dimer of active dimers [6]. Because glutamate-bound LBDs are closed, ligand binding is thought to pull open the gate of the attached

ion channel pore. Alternatively, binding can drive separate conformational changes that result in an unresponsive, desensitized receptor. A paradox of AMPA receptor activation is how the dynamics of both binding and unbinding of glutamate are fast enough in order to allow both the onset and termination of activity to be rapid, yet activation is robust and selective.

Over 100 high-resolution structures of genetically isolated iGluR LBDs have been determined, in complex with agonists and competitive antagonists, as well as in the apo state. These structures reveal common modes of ligand binding within the cleft [6]. Computational analyses have shed light on both LBD and ligand dynamics and energetics [7–18]. Neither the structural nor computational studies, have comprehensively shown how glutamate finds its way into the binding site or how large-scale conformational changes in the LBD are coupled to glutamate binding. For example, does glutamate diffuse directly from bulk solvent into its binding pocket, only contacting the protein randomly, or does glutamate follow distinct pathways on the surface of the LBD to find its way into its pocket? If the latter, what structural features make up the pathways, and what is the nature of the protein-ligand interactions therein? In this study, long-timescale molecular dynamics simulations (totaling nearly 50 μs) and free energy calculations of the GluA2 receptor suggest that strategically-positioned flexible sidechains on the surface of the LBD metastably interact with glutamate to help funnel it into its recessed binding pocket, where it adopts two possible poses. These results were used to guide the design of a panel of LBD mutants that were tested using electrophysiological recordings. Elimination of the transient binding sites was found to slow both activation and deactivation of the receptor. Taken together, these results suggest that glutamate, and perhaps other iGluR ligands as well, binds via distinct pathways on the surface of the LBD, and disruption of these pathways significantly impacts the functional kinetics of the receptor.

4.2 Results and Discussion

4.2.1 Glutamate binds via preferred pathways and metastable interactions

In order to examine the processes by which a glutamate ligand either associates with or dissociates from the GluA2 AMPA receptor LBD, we performed unbiased all-atom MD simulations with explicit solvent using special-purpose hardware [19] to generate 21 trajectories with an aggregate of 49.1 $/mu$s. The four LBDs of an iGluR are arranged as a dimer of dimers; we simulated glutamate binding in both LBD dimers and monomers. The binding of glutamate is thought to involve: ligand entry into an open, solvent-exposed binding cleft, ligand docking to R485 in Lobe 1 of the LBD, then large- scale conformational changes that close the cleft, securing the ligand [20]. After glutamate has docked, cleft closure proceeds when the ligand, attached to R485, also attaches to E705 in Lobe 2 [21].The ligand forms both of these Lobe 1 and Lobe 2 interactions in all of the simulated binding trajectories.

In our simulation systems (see **Table 4.1 Table 4.2**), glutamate ligand molecules were initially placed at random positions and orientations in bulk solvent, similar to the approach of Dror et al. [22], at least 8 Å away from any non-water molecules. The effective ligand concentrations ranged from 3.9 mM (single ligand in a monomer system) to 71 mM (20 ligands in a dimer system). These concentrations are somewhat higher than estimates of the peak glutamate concentration during synaptic transmission (1-10 mM) [23, 24], but our goal was to maximize the likelihood of observing multiple independent binding events within our allocated simulation timeframe. Despite the high concentrations, we did not observe any artifactual ligand- ligand or protein-ligand interactions. In all but one of the binding trajectories (see 4.7), the two negatively charged carboxyl groups of glutamate are exploited by positively-charged side chains at the periphery of the binding cleft. Strikingly, glutamate was passed from one residue to the next by a select set of residues that "funnel" the neurotransmitter, via a series of metastable interactions, into its binding site.

In one such trajectory (**Fig. 4-1**, Movie S1), the ligand first contacts the LBD at

R453. From there, its γ-carboxyl group swings toward Lobe 2 to interact with helix F residues E657, R660, and R661. Mutations in this region, a prominent metastable interaction site (occurring in three out of six binding simulations), alter AMPA receptor function [25]. R660 is conserved between AMPA and NMDA receptors; in kainate receptors, R660 and R661 are replaced by lysine residues (**Fig. 4-8**a). The guanidinium group of R485 rotates dramatically out of the binding pocket to contact the ligand. This surprising conformational flexibility occurs via rotations of $-106°$, $-127°$, $-136°$, and $-111°$ around the R485 sidechain's χ_1, χ_1, χ_3, and χ_4 torsion angles, respectively (**Fig. 4-1**, e-i). The ligand remains tethered to R485 as it leaves the metastable site, being pulled into the binding cleft. Finally, the cleft closes to secure the bound ligand in the crystallographic pose. Details of individual trajectories are provided in 4.7.

The binding pathways in the dimer and monomer simulations are very similar. Given the functional independence of binding processes and the location of the pathways at the periphery of the LBD tetramer (**Fig. 4-8**b), we expect that the pathways we observe are the same in full-length, tetrameric forms.

In all trajectories, K730, in the hinge region between Lobes 1 and 2, shows substantial conformational flexibility and alternates between forming salt bridge interactions with E705 in Lobe 2 and D728 in the hinge. From our simulations, it is unclear how much these interactions contribute to "locking" the LBD closed, as first proposed by Armstrong and Gouaux [21]. Water molecules were observed to occupy approximately the same positions in the cleft as seen in crystal structures [21].

4.2.2 Potential of mean force for binding and unbinding

To understand the energetics of funneling in each of the pathways, we computed a three- dimensional free energy landscape, or potential of mean force (PMF), from our binding and unbinding trajectories. The PMF, shown in **Fig. 4-2**a, indicates three possible ligand binding pathways. Contouring the PMF at lower energies revealed sites of metastable protein-ligand interactions (**Fig. 4-2**b-d). Situated inside the binding pocket, site 0 is the global free energy minimum, set to 0 kcal/mol. The

next most stable sites, sites 1, 2, and 3, are local minima with free energies of 0.29, 0.83, and 0.75 kcal/mol, respectively, located at positions where the ligand forms metastable interactions with the LBD. Sites in bulk solvent, on the other hand, have free energies of about 3.12 kcal/mol. The ligand traverses pathway 1 in about half of the trajectories, pausing at site 1 prior to binding at site 0. In pathway 2, the ligand interacts with sites 2 and 1 prior to binding. In pathway 3, the ligand transitions from site 3 directly to site 0. Residue-ligand interactions play a role in passing the ligand from each site to the next. In particular, helix F residues (E657, R660, and R661) are involved in interactions at site 1 and 2 as well as transitions from site 2 to site 1 and from site 1 to site 0. K449 is involved in shuttling the ligand into the binding pocket from site 3. A list of the residues that interact with the ligand at each site is provided in the legend for **Fig. 4-2**a. An error analysis of the PMF is provided in **Fig. 4-9**.

The association rate constant calculated from our binding trajectories, $k_{\text{on}} = 1.4 \times 10^7$ $\text{M}^{-1}\text{s}^{-1}$, using the approach of Dror *et al.* [22], agrees closely with the experimentally measured value of 1.6×10^7 $\text{M}^{-1}\text{s}^{-1}$ for the GluA4 LBD [26] (**Table 4.2**). To our knowledge, no experimentally measured k_{on} for the GluA2 LBD has been reported, but the submillisecond activation lag at 10 mM glutamate reported here and elsewhere is consistent with this value.

4.2.3 Glutamate can bind in an inverted pose

Two of the binding trajectories resulted in glutamate binding into the crystallographically observed pose [21], in which the ligand's α-carboxyl group is anchored by R485 in Lobe 1 while the γ-carboxyl group is stabilized by the backbone amides of S654 in Lobe 2 (**Fig. 4-3**a). In the crystal structure, the γ- carboxyl group also interacts with the backbone amide of T655 in Lobe 2, but this interaction was not observed in our binding trajectories; further closure of the LBD would be required for this interaction. The other four trajectories unexpectedly resulted in glutamate binding into an "inverted" pose, in which the γ-carboxyl group binds to Lobe 1, and the α-carboxyl group binds to Lobe 2 (**Fig. 4-3**b and **Fig. 4-10**, Movie S2). The lack of preference for

binding the α- or γ-carboxylates of glutamate at metastable sites presents a paradox. How is glutamate delivered efficiently into the pose seen in crystallographic experiments? Indeed, in several of the simulations, glutamate adopts an inverted pose. However, in one simulation (**Fig. 4-11**, Movie S3, T_{dim2} in **Table 4.1**), glutamate rotates from the crystallographic pose to the inverted pose, suggesting interconversion between poses is possible. In this case, the binding cleft was not fully closed. Free energy landscapes governing cleft closure have been described using the two-dimensional order parameter (ξ_1, ξ_2) (**Fig. 4-3**c)([10, 14, 17, 27]). The extent of cleft closure for LBD conformations with glutamate bound in the inverted pose is not as great as that seen in crystal structures or in T_{dim1} (**Fig. 4-3**d, Movie S1). It is possible that an extension of T_{dim2} could have resulted in the ligand returning to the crystallographic pose. In support of this notion, MD simulations of drug binding to G-protein-coupled receptors showed that a drug initially bound in a non-crystallographic pose eventually converted to the crystallographic pose [22]. Our measurements do not indicate how stable the non-crystallographic, inverted pose is, but this binding mode could cause glutamate to act as a partial agonist with a smaller degree of cleft closure. Individual AMPA receptor activations show substantial sublevel activity, connected to LBD occupancy [28]. Stochastic fluctuations in the single channel current could additionally reflect dynamic conversions between crystallographic and non-crystallographic ligand poses.

4.2.4 Unbinding pathways mirror binding pathways

We also simulated ligand dissociation trajectories, which we initiated from either crystal structure-like configurations in which the LBD is fully closed with glutamate in the crystallographic pose (PDB: 1FTJ) or ligand-docked configurations in which the LBD is semi- closed or open (**Table 4.1**). The latter trajectories were continuations of prior ligand-association trajectories. In general, for both the LBD dimer and monomer systems, binding and unbinding pathways appear to be the reverse of each other, with free energy barriers being traversed in opposite order. The ligand exited from the ξ_1 side of the binding cleft in all five dissociation events involving the dimer.

For the monomer, the ligand exited from the ξ_1 side in four dissociation events and from the ξ_2 side in two events.

4.2.5 Disrupting binding pathways selectively slows both activation and deactivation

The simulations suggest that clusters of charged residues along the binding pathways should interact with glutamate during both association and dissociation. To test this hypothesis, we designed a set of single, double and triple mutants, removing or reversing the polarity of charged sidechains at each metastable site in turn (**Fig. 4-4**a). We hypothesized that these mutations should alter receptor activation, deactivation, and perhaps recovery from desensitization. Given that the metastable sites are formed by flexible residues lacking interactions with other parts of the receptor, we expected little effect of the mutations on downstream gating conformational changes. Additionally, we expected that entry into desensitization, which likely involves neither binding nor conformational changes of the individual LBDs, should be unaffected

Association is the first step in the sequence of events from resting to activated states. Therefore, we applied long (200 ms) pulses of 10 mM glutamate, to measure the activation as a proxy for glutamate association, with only minimal contamination from the ∼50-fold slower desensitization process. Activation of wild-type (WT) GluA2 by 10 mM glutamate is very fast, with a 10-90% rise time of 200 ± 30 μs. The rise time is probably slowed by the rate of solution exchange onto the excised patch. Strikingly, a substantial 3-fold increase in the 10-90% rise time during activation, t_{rise}, was observed for R453D and RDK-AAA (t_{rise} = 645 ± 60 and 760 ± 70 μs, respectively; $P < 0.005$ vs. WT GluA2, t-test, n = 3-7; **Fig. 4-4**b,c; **Table 4.3**).

We next investigated receptor deactivation, a process limited by the glutamate unbinding rate, by applying short (1 ms) pulses of 10 mM glutamate. WT GluA2 deactivates rapidly, with a deactivation time constant, τ_{deact}, of 1.5 ± 0.2 ms. Mutating residues involved in metastable site 1, we observed a robust increase in τ_{deact} for R453D (τ_{deact} = 3.6 ± 0.1 ms; **Table 4.3**) and the triple alanine mutant, R453A-D456A-K458A (denoted RDK-AAA; τ_{deact} = 3.7 ± 0.1 ms), compared with WT

GluA2 (P = 0.003, t-test, n = 3-5) (**Fig. 4-4**d).

Single substitutions on helix F (R660 and R661; sites 1 and 2) had little effect on activation or deactivation, but the triple mutant E657A-R660A-R661A (denoted ERR-AAA, **Fig. 4-5**a) exhibited the same profile of slower activation (t_{rise} = 550 ± 80 μs; **Fig. 4-5**b; **Table 4.3**) and deactivation (τ_{deact} = 3.8 ± 0.4 ms; **Fig. 4-5**c) as the RDK-AAA mutant. A comparison of EC_{50} values for WT GluA2 (EC_{50} = 330 ± 50 μM) and ERR-AAA (EC_{50} = 410 ± 30 μM), revealed that the apparent affinity of the ERR-AAA mutant for glutamate is similar to that of WT (P = 0.5; t-test, n = 3) (**Fig. 4-12**d). This result is expected if mutations remove metastable interactions but do not affect the initial or final conformational states, revealing a distinct mechanism to previous reports of LBD mutations that change binding [25, 29].

We also tested the RDK-AAA mutant with 50 mM glutamate, to ensure that ligand diffusion was not a factor in slowed activation. Both activation and deactivation remained substantially slower than in WT GluA2 (t_{rise} = 660 ± 70 μs; P = 0.01 vs. WT GluA2, t-test, n = 3; τ_{deact} = 3.7 ± 0.1 ms; P = 0.0003 vs. WT GluA2, t-test , n = 3; **Fig. 4-12**, a, b, **Table 4.3**). This result strongly suggests that the binding kinetics are principally altered by mutations at these metastable sites.

In order to determine how the ERR-AAA mutant perturbs glutamate binding, we performed unbiased simulations involving this mutant LBD (Movie S4, **Table 4.4**). PMF calculations indicate that this LBD lacks metastable interactions along Pathways 1 and 2, proximal to helix F (**Fig. 4-5**e). The dominant binding pathway, which does not involve helix F interactions, is similar to Pathway 3 in the WT LBD.

As expected for mutations that perturb only binding kinetics, desensitization was weakly perturbed by mutations to the ligand binding pathways. The time constant of entry to desensitization, τ_{desen}, for RDK-AAA, ERR-AAA, and single point mutants, were similar to WT GluA2 (**Fig. 4-13**a,b, and **Fig. 4-12**c, and **Table 4.3**). For most mutants, the time constant of recovery from desensitization was similar to WT GluA2 (60 ± 5 ms), but the R453D and RDK-AAA mutants recovered about twice as fast (P=0.01 and P=0.02, t-test, n = 3; **Fig. 4-13**c,d). These results support the idea that entry to and recovery from desensitization occur with the clamshell closed

and glutamate stably bound, even though glutamate must eventually unbind during recovery.

The second set of mutants that we generated targeted pathway 3. The D447A-K449A double mutant in Lobe 1 (DK-AA) slowed activation ($t_{\text{rise}} = 410 \pm 40$ μs ; P = 0.01, vs. WT GluA2, t-test, n = 4-6, **Fig. 4-6** a, b, **Table 4.3**) compared with WT GluA2. The very poor expression of this mutant precluded a robust measurement of deactivation ($\tau_{\text{deact}} = 3 \pm 0.3$ ms; P = 0.1 vs. WT GluA2, t-test, $n = 3$; **Fig. 4-4**C). The R684A-E688A double mutant in Lobe 2 (RE-AA), which targeted site 3, slowed both t_{rise} and τ_{deact} to a greater extent ($t_{\text{rise}} = 590 \pm 50$ μs and $\tau_{\text{deact}} = 4.9 \pm 0.7$ ms; $P < 0.05$, t-test, $n = 3$, **Fig. 4-6**b,c). For the RE-AA mutant, we also observed an apparent acceleration in of recovery from desensitization ($\tau_{\text{rec}} = 13 \pm 1$ ms, $n = 3$; **Fig. 4-14**) compared with WT GluA2, although again, poor expression made proper estimation of the recovery time constant difficult.

4.2.6 Off-pathway mutants do not perturb binding

Substantial charge swap mutations could have non-specific effects on binding such as altering binding domain structure, electrostatic profiles or intersubunit interactions. Therefore, as negative controls, we generated complementary, off-pathway mutations that our simulations predicted did not interact with glutamate during binding and unbinding. Two triple alanine mutations were generated: K409A-K410A-E422A (denoted KKE-AAA) in Lobe 1 and R715A- K716A-D769A (denoted RKD-AAA) in Lobe 2. These off-pathway mutants had smaller effects on kinetics, ($t_{\text{rise}} = 100 \pm 5$ μs, $\tau_{\text{deact}} = 2.4 \pm 0.1$ ms for RKD; $t_{\text{rise}} = 310 \pm 10$ μs, $\tau_{\text{deact}} = 1.1 \pm 0.1$ ms for KKE; **Fig. 4-6**d-f). Their kinetics conformed to previously published mutants that influence closed-cleft stability. Quite distinct from the metastable binding sites presented here, for RKD-AAA slower deactivation was accompanied by faster activation (**Table 4.3**; indicative of an increase in affinity). For KKE-AAA, the results were inverted, with faster deactivation and slower activation slower (consistent with reduced glutamate affinity).

For most of the mutants tested, the decay rate is slowed more profoundly than

the opening rate. The decay time constant varies linearly as a function of the activation time, and changes in the decay are roughly 8-fold larger than that of activation (**Fig. 4-7**c). Simulated currents using previously published kinetic models of AMPA receptor activation [29] show that this relationship is precisely what is expected from slowing glutamate association and dissociation rates by equal amounts. Moreover, these calculations were performed with realistic concentration jump profiles [30], indicating the physically plausible solution exchange rate of about 300 μs would be expected to yield a very similar, roughly linear relationship between activation and deactivation times (**Fig. 4-7**). Notably, the off-pathway mutants KKE-AAA and RKD- AAA (**Fig. 4-6**d-f) lie away from the linear relation between activation and decay time (**Fig. 4-7**c).

Taken together, these results strongly support the hypothesis that these metastable sites predicted *in silico* form preferential pathways to guide glutamate in and out of its binding site.

4.3 Conclusion

By combining long unbiased simulations with direct measurements of receptor kinetics, we could show a key role for charged residues in facilitating fast neurotransmitter access to a deep binding site. These studies show that neurotransmitter binding is a directed process for which kinetics have been optimized (presumably by evolution) without altering overall ligand affinity. Previous work has shown that electrodiffusion of glutamate in the synaptic cleft speeds up neurotransmission [31]. Our experiments reveal a strikingly elaborate management of ligand transport by AMPA receptors, whereby flexible positive charges ensure that glutamate binding reactions are fast. The existence of these pathways is surprising, and the fact that they alter the kinetics of receptor activity indicates that the molecular mechanisms that determine the action of neurotransmitters at receptors are more complex than previously thought. Given that electrostatic interactions are also important for coordination in other neurotransmitter binding sites [32], these principles of ligand funnelling may be general.

4.4 Materials and Methods

4.4.1 Simulation system preparation.

The initial atomic models for both the monomer and dimer systems were constructed from the crystal structure of the GluA2 ligand-binding core (S1S2) in complex with glutamate (PDB ID: 1FTJ). Missing amino acid residues were built using the Modloop server [33], and missing sidechains were built using SCWRL4 [34]. Crystallographic waters in the ligand-binding cleft were included. The monomer system, which contained a total of 47,227 atoms, was solvated with 14,369 water molecules and neutralized by adding Na^+ and Cl^- ions to the bulk solution until the salt concentration was 150 mM NaCl. Periodic boundary conditions were imposed on an orthorhombic cell with approximate dimensions 88 Å × 68 Å × 78 Å. The dimer system, which contained a total of 56,217 atoms, was solvated with 15,951 water molecules and neutralized to maintain 150 mM NaCl. Periodic boundary conditions were imposed on an orthorhombic unit cell of approximate dimensions 96 Å × 78 Å × 78 Å. The system was energy minimized and equilibrated using constant pressure and temperature (NPT) conditions at 1 atm and 300 K with a timestep of 2 fs. The all-atom potential energy function PARAM27 for proteins [35, 36] and the TIP3P potential energy function for water [37] were used. Electrostatic interactions were computed using the particle mesh Ewald (PME) algorithm and short-range, non-bonded interactions were truncated to 12 Å. The initial protein configuration of the system was relaxed with Langevin dynamics in the presence of harmonic restraints at constant volume for 30 ps before the barostat was switched on at 1 atm for a further 60 ps of simulation in NPT conditions. The cell dimensions were allowed to vary for 2 ns in NPT conditions before reaching the final box size. A 4 ns pre-production run in constant volume and temperature (NVT) conditions was carried out from which five, ligand-bound, starting coordinates for the monomer, and one, ligand-bound, starting coordinates for the dimer were selected for long-timescale simulation. The pre-production run was performed using NAMD 2.9 [38], while minimization and equilibration procedures were done in CHARMM [39].

4.4.2 System preparation for the ERR-AAA simulations.

Starting coordinates were selected from trajectory T_{mon2}, which involved the monomer system containing 10 ligands. Residues E657, R660, and R661 were mutated to alanine in CHARMM by deleting the side chain atoms and replacing with methyl groups. The salt concentration was adjusted to maintain 150 mM NaCl. The mutant system containing 47,080 atoms was energy minimized and briefly equilibrated in a 2 ns pre-production run in constant NVT conditions.

4.4.3 Simulations with increased ligand concentration.

For the LBD monomer, 10 ligands were added at arbitrary positions in bulk solvent, each greater than 20 Å from the binding pocket, to a previously prepared system containing an LBD in an open conformation. Salt concentration was adjusted for ligand charge to maintain 150 mM NaCl. A 2 ns pre-production run was carried out in constant NVT conditions. For the LBD dimer, 20 ligands were added at arbitrary positions in bulk solvent greater than 20 Å from the binding pocket to a previously prepared system containing open conformations of the LBDs. Salt concentration was adjusted for ligand charge to maintain 150 mM NaCl. A 2 ns pre-production run was carried out in constant NVT conditions.

4.4.4 MD simulations.

All production runs used the NPT ensemble at 1 atm and 300K. Bond lengths for hydrogen atoms were constrained using the M-SHAKE algorithm [40]. An r-RESPA integrator [41] was used with a timestep of 2 fs; long-range electrostatics were computed every 6 fs. Long-range electrostatics interactions were calculated using the k-space Gaussian split Ewald method [42] with a 64 Å × 64 Å × 64 Å grid, $\sigma = 2.02$ Å, $\sigma_s = 1.29$ Å. Short-range interactions including van der Waals and short-range electrostatics were truncated at 9 Å. To prevent overall rotational and translational motion of the protein, positional harmonic restraints were applied on the backbone atoms of residues 426-428 (residues 37-39 in 1FTJ), residues 474-476 (residues 85-87 in 1FTJ), and residues 490-492 (residues 101-103 in 1FTJ) with a force constant of 0.3

kcal mol^1 Å^{-2}. All productions simulations were carried out on the special purpose Anton machine at the Pittsburgh Supercomputing Center [19]. Production simulations for the ERR-AAA mutant protein were carried out on the Anton2 machine at the PSC [43]. Simulations on Anton2 were carried out as they were on Anton, except using a temperature of 310K, which is the default for Anton2. A total of 21 trajectories were generated for an aggregate simulation time of 49.1 μs: 11.8 μs involved WT dimers, 36.1 μs involved WT monomers, and 1.2 μs involved the ERR-AAA monomer.

4.4.5 Ligand binding PMF.

The trajectories containing the monomeric LBD with 10 ligands ($T_{mon2,3}$) was sampled at 0.12 ns intervals. Cartesian coordinates for the ligand's non-hydrogen atoms, $\vec{r}$, were measured and used as three-dimensional order parameters to describe the states along the binding pathway. Each frame in the trajectory was aligned with respect to the backbone atoms of the LBD. The density of atomic positions, $\rho(\vec{r})$, was computed using a hard sphere van der Waals approximation onto a discretized grid with a spacing of 0.5Å $\times$ 0.5 Å $\times$ 0.5 Å and subsequently weighted to produce the free energy maps using the standard Boltzmann re- weighting scheme, i.e., $W(\vec{r}) = -k_B T \log[\rho(\vec{r})]$. The statistical uncertainty in the PMF was determined using the approach of block averaging [44] (**Fig. 4-9**). The trajectory was subdivided into 10 blocks, and a PMF was calculated for each block. The standard deviation in the 10 PMFs was calculated. Using 5âĂŞ15 blocks all gave qualitatively similar results.

4.4.6 Calculation of k_{on} from molecular simulations.

$k_{\text{on}} = \frac{N_b}{\sum_i \frac{t_i [L_i]}{s_i}}$, where N_b is the total number of binding events; t_i is the time the ligand spends in bulk solvent, $[L_i]$ is the free ligand concentration, s_i is the number of protein subunits, and i is summed over all simulation systems. k_{on} was calculated irrespective of the bound pose of the ligand and suggests glutamate binding is not diffusion controlled since it is 100 times slower than the association rates of typical diffusion-limited binding processes (e.g., $k_{\text{on}} \sim 10^9 - 10^{10}$ M^{-1} s^{-1}) [45, 46]. See **Ta-**

ble 4.2. Electrostatic steering may play a role in determining the overall association rates of the ligand to the binding site [47].

4.4.7 Molecular biology.

Point mutations were introduced by overlap PCR and confirmed by double-stranded sequencing. Numbering refers to the mature polypeptide chain.

4.4.8 Electrophysiology.

WT and mutant glutamate receptors were overexpressed in HEK293 cells using calcium phosphate transfection. The external solution contained: 150 mM NaCl, 0.1 mM $MgCl_2$, 0.1 mM $CaCl_2$, and 5 mM HEPES, titrated to pH 7.3 with NaOH, to which we added drugs as required. Drugs were obtained from Ascent Scientific and Sigma. The pipette solution contained: 115 mM NaCl, 10 mM NaF, 0.5 mM $CaCl_2$, 1 mM $MgCl_2$, 5 mM Na_4BAPTA, 5 mM HEPES and 10 mM Na_2ATP (pH 7.3).The sampling rate was 10 kHz (100 μs time step) and during acquisition the data were filtered at 5 kHz (10-90% rise time 66 mus). We applied ligands to outside out patches via a piezo-driven fast perfusion system. Typical 10%-90% solution exchange times were faster than 300 μs, as measured from junction potentials at the open tip of the patch pipette. Simulations using realistic concentration jumps were done as described [30] using a suite of PYTHON scripts (https://github.com/aplested/aligator). The effective decay constant was back-calculated from the 90âĂŞ10% decay time assuming a single exponential decay to facilitate comparison. These simulations showed that a solution exchange of 300 μs reproduces all the observed features of glutamate- activated AMPA receptor currents, including the approximately 200 μs rise time of wild-type channels **Fig. 4-7**.

4.4.9 Functional data analysis.

To measure deactivation and desensitization decay constants, we fitted currents with a single exponential function. The rise times were measured from the 10-90% crossing times of a sigmoid function fitted from the baseline to the peak current. We mea-

sured concentration-response curves for WT and the mutant receptor EKK-AAA. We obtained the EC_{50} and maximum extent of activation relative to glutamate from fits to the Hill equation

$$\frac{I}{I_{\max}} = \frac{[\mathrm{A}]^n}{[\mathrm{A}]^n + [\mathrm{EC}_{50}]^n}$$

where n is the Hill coefficient, $I_{\max}$ is the maximum response, and [A] is the agonist concentration.

To measure recovery from desensitization, we used a two-pulse protocol with a variable interpulse interval. Recovery data were fitted by a Hodgkin- Huxley-type function [48]

$$N = N_0 + (1 - N_0)[1 - \exp(-k_{\mathrm{rec}}t)]^n$$

where N is the active fraction of receptors at time t following the first pulse, N_0 is the active fraction at the end of the conditioning pulse, and k_{rec} is the rate of recovery.

4.5 Author Contributions

A.Y., H.S., A.J.R.P., and A.Y.L. designed research. A.Y. and H.S. performed research. A.Y. and A.J.R.P. contributed new reagents/computational tools. A.Y., H.S., A.J.R.P., and A.Y.L. wrote the paper.

4.6 Acknowledgments

We thank Jelena Baranovic for comments on the manuscript. Anton computer time was provided by the Pittsburgh Supercomputing Center (PSC) through Grant R01GM116961 from the National Institutes of Health. The Anton machine at PSC was generously made available by D.E. Shaw Research. This work was supported by the National Institutes of Health grant R01GM094495 (to A.Y.L.), the ERC grant "GluActive" (647895) and the DFG Cluster of Excellence "NeuroCure" (EXC-257) (to A.J.R.P.). A.J.R.P. is a Heisenberg Professor of the DFG (PL619-3.1). H.S. was the recipient of a fellowship from the Human Frontier Science Foundation.

4.7 Additional trajectory descriptions

Trajectory T_{dim2}

In T_{dim2} , glutamate contacts K449 in Lobe 1 on the anterior surface of the LBD and then moves towards the binding cleft, sandwiched between K449 and R453. The α-carboxyl group transitions away from K449 to contact R485 and temporarily adopts the crystallographic pose (**Fig. 4-3**A). Contacts between Lobe 2 and the ligand are then severed. Subsequently, the entire ligand rotates such that its γ-carboxyl group now contacts R485, adopting the inverted pose (Fig. 3B); the α-carboxyl group interacts with S654 in Lobe 2; the amide is coordinated by P478, T480, E705, and a water molecule. The events of T_{dim2} (**Fig. 4-9** and Movie S2) suggest interconversion between poses is possible, at least before the cleft completely closes. It is possible that an extension of T_{dim2} could result in the ligand returning to the crystallographic pose. In support of this notion, MD simulations of drug binding to G-protein-coupled receptors showed that a drug initially bound in a non-crystallographic pose, it eventually converted to the crystallographic pose [22]. Interestingly, the extent of cleft closure on the inverted glutamate is not as great as that seen in T_{dim1}. It is unknown how stable the inverted pose is, but this binding mode could effectively cause glutamate to act as a partial agonist due to a lesser extent of cleft closure.

Trajectories T_{dim3} and T_{dim4}

The trajectories T_{dim3} and T_{dim4} also result in an inverted pose. T_{dim3} resembles T_{dim2} up to the point at which the ligand's Îş-carboxyl group contacts R453 and the α-carboxyl group contacts K449. In T_{dim3}, however, instead of the α-carboxyl group contacting R485, which would result in the crystallographic pose, the ligand's amide group transitions into the binding cleft to contact E402 on the ξ_2 side. The γ-carboxyl group then switches interaction partners from R453 to R485 and settles into its bound pose. In T_{dim4}, the ligand diffuses onto the Lobe 2 surface of the binding cleft and contacts R485 with its γ-carboxyl group. T_{dim4} involves no significant metastable interactions before the ligand assumes its bound pose.

Trajectory T_{mon1}

For one of the two ligand-association events that were observed with an LBD monomer, trajectory T_{mon1} , the ligand diffused in bulk solvent for 2.95 μs before binding to the LBD, presumably because this system contained only one glutamate ligand (effective concentration of 3.9 mM). T_{mon1} initially resembles T_{dim1} except that glutamate binds in the inverted pose (**Fig. 4-10** and Movie S3). In T_{mon1}, the ligand initiates contact with the LBD by forming a metastable interaction with E657, R660, and R661 on helix F via its α-carboxyl group. Its γ-carboxyl group subsequently binds to the guanidinium group of R485, which, as in T_{dim1}, has transiently rotated out of the binding pocket to form the interaction. Helix F releases its hold on the ligand, which then binds into the binding cleft via the same interactions described for T_{dim2}.

Trajectory T_{mon2}

The other association event involving an LBD monomer took place within a system containing 10 glutamate ligands (effective concentration of 38.9 mM). The binding pathway very closely resembles T_{mon1}, with the ligand forming a metastable interaction with R660 and R661 on helix F before binding to R485 and docking in the inverted pose.

Fig. 4-1: Dynamics of glutamate binding. Time points (lower left of each panel) are relative to the start of Movie S1. This trajectory corresponds to the LBD dimer system T_{dim1} (see **Table 4.1**); only the LBD that binds glutamate is shown for clarity. (a) Prior to ligand entry to the binding pocket, the LBD is open; (ξ_1, ξ_2) = (12.3, 12.2 Å). The ligand's γ-carboxyl contacts R453 on Lobe 1. (b) Close-up view of (a). (c) Glutamate slips into the binding cleft. (d) The ligand contacts R661 on Lobe 2. (e) A metastable interaction forms across Lobes 1 and 2. The ligand's γ-carboxyl contacts E657, R660, and R661 on helix F and the ligand's α-carboxyl contacts R485. R485 flickers out of the binding pocket to interact with the ligand. (f-h) R485 relaxes towards the binding pocket. The metastable interaction at the ligand's α- and γ-carboxyl between Lobes 1 and 2 persists. (i) The ligand shifts into the binding pocket, with its α-carboxyl contacting R485. Lobe 2 interactions with helix F are broken. In the pocket, the ligand's amide is coordinated by P478 and L480. Cleft closure is initiated once helix F undergoes a backwards tilt to form a pocket for the ligand's γ-carboxylate. (j) Glutamate adopts the crystallographic conformation. (k) The ligand's amide contacts E705 on Lobe 2. (l) Expanded view of (k). The LBD closes around the ligand in the crystallographic conformation: (ξ_1, ξ_2) = (11.8, 10.8 Å). See also **Fig. 4-8**, Movie S1, and **Table 4.1**.

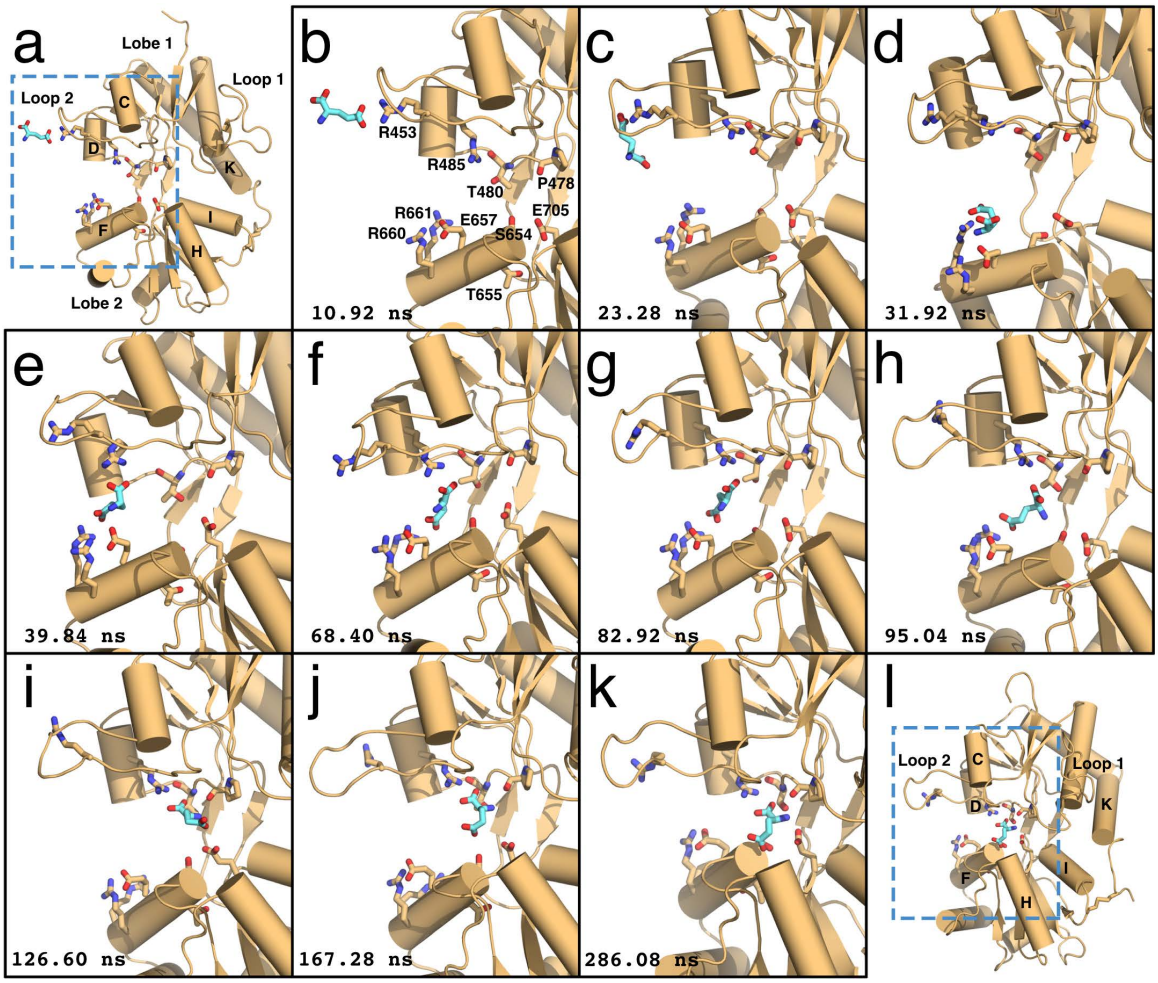
a
Lobe 1
Loop 1
Loop 2
C
D
K
F
I
H
Lobe 2
b
R453
R485
T480
P478
R661
E657
E705
R660
S654
T655
10.92 ns
c
23.28 ns
d
31.92 ns
e
39.84 ns
f
68.40 ns
g
82.92 ns
h
95.04 ns
i
126.60 ns
j
167.28 ns
k
286.08 ns
l
Loop 2
C
Loop 1
D
K
F
I
H

Fig. 4-1

Fig. 4-2: Glutamate binding pathways and metastable binding sites. (a) The PMF calculated from the ligand density using a hard-sphere van der Waals approximation on a grid spacing of 0.5 Å along the x, y, and z axes contoured to 1.89 kcal mol^{-1}. The primary binding pathways for glutamate are depicted by arrows. (b) The PMF, contoured to 1.16 kcal mol^{-1}, shows the metastable binding sites, sites 1-3. Site 0 is the site of stable binding. Site 1 is shared by pathways 1 and 2, site 2 is encountered in pathway 2, and site 3 is encountered in pathway 3. Site 1 spans the two lobes, involving R453 in Lobe 1, and E657, R660, and R661 in Lobe 2. Site 2 involves E657, R660, and R661 in Lobe 2. Site 3 involves R675 and R684 in Lobe 2. The residues involved in Site 0 are shown in panel **Fig. 4-3**a. Error analysis is provided in **Fig. 4-9**. (c) The PMF, contoured at 0.32 kcal mol^{-1}, shows the global free energy minimum. This minimum overlaps well with the ligand density derived from crystal structures of the glutamate-bound complex (e.g., PDB ID: 1FTJ). (d) The one-dimensional representation of the WT GluA2 ligand-binding PMF was obtained by first computing the three-dimensional PMF (panels a-c). Values of the three-dimensional PMF were indexed increasing in the z, y, then x directions, from ($x_{\min}$, $y_{\min}$, $z_{\min}$), to produce the one-dimensional representation. The positions of Sites 0-3 are indicated. Site 0 is the global free energy minimum and is set to 0 kcal/mol; Sites 1-3 form local minima with free energies of 0.29, 0.83, and 0.75 kcal/mol, respectively. See also **Fig. 4-9**.

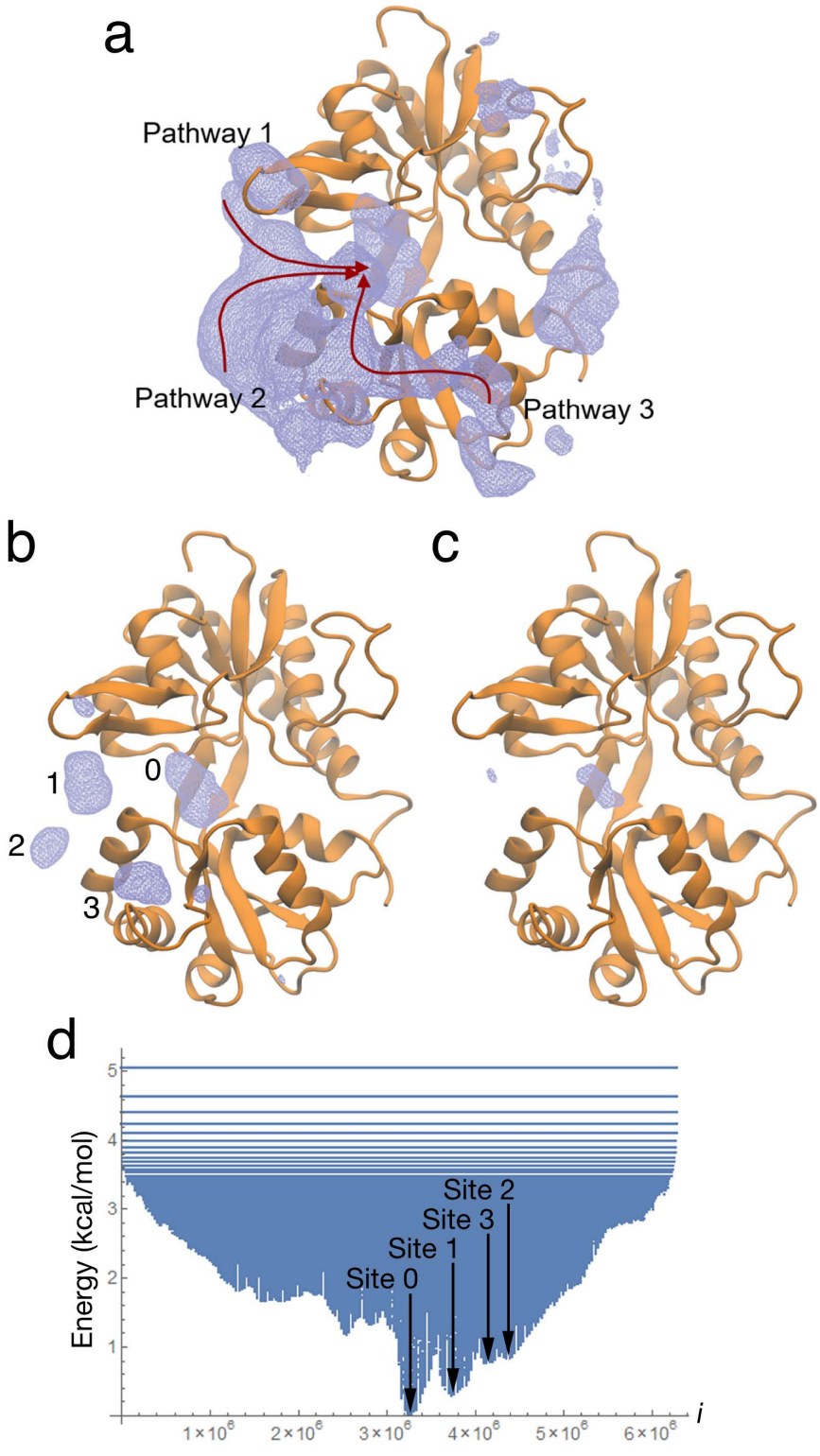
a
Pathway 1
Pathway 2
Pathway 3
b
0
1
2
3
c
d
Energy (kcal/mol)
Site 0
Site 1
Site 2
Site 3
i
1×10^6
2×10^6
3×10^6
4×10^6
5×10^6
6×10^6

Fig. 4-2

Fig. 4-3 Conformations of bound glutamate and the LBD. (a) The bound ligand conformation, similar to that observed in PDB ID 1FTJ. The ligand's α-carboxylate contacts R485, and its γ-carboxylate contacts the amide backbone of T655. The ligand's amide group is coordinated to the sidechains of T480 and E705 and to the backbone of P478. (b) The inverted conformation of the ligand. The ligand's α-carboxylate contacts the backbone amide of S654, and its γ-carboxylate contacts R485. The ligand's amide group is coordinated to E705. P478 has moved upwards to accommodate a water molecule that also contacts the ligand's amide group. (c) The two-dimensional order parameter (ξ_1, ξ_2) used to characterize large-scale conformational transitions in the GluA2 LBD. ξ_1 and ξ_2 each indicate the distance between the centers-of-mass of the clusters of atoms shown in blue and green, respectively. (d) (ξ_1, ξ_2) measures the degree of cleft closure for the LBD in apo (blue) and ligand-bound conformations. The ligand occupies either the crystallographic (yellow) or inverted (green) poses. Each point represents a snapshot taken every 120 ps from simulations of the monomer system at 3.9 mM glutamate concentration. The marginal histograms indicate distribution densities. See also **Fig. 4-10** and **Fig. 4-11**, and Movies S2 and S3

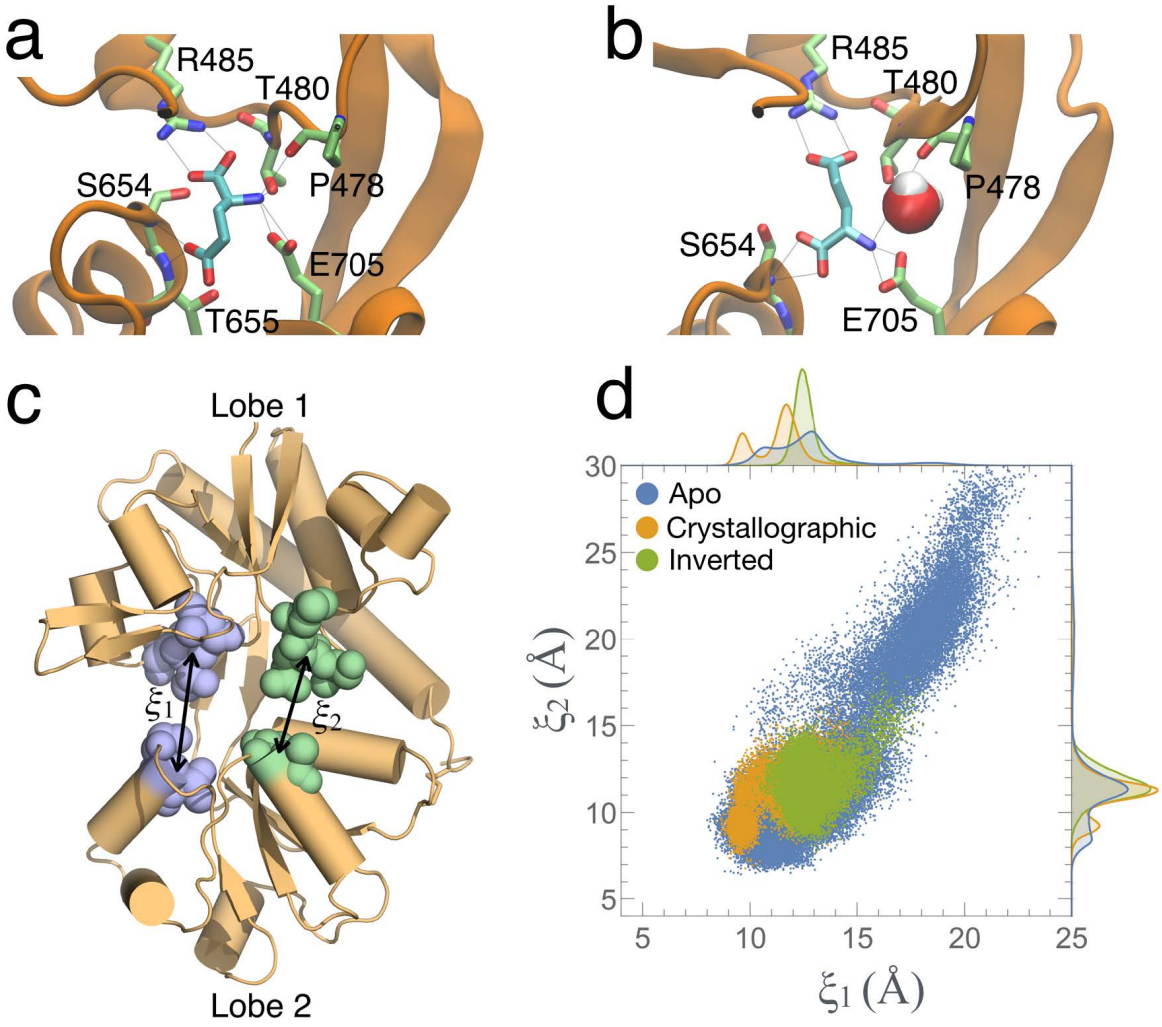
a
R485
T480
S654
P478
E705
T655
b
R485
T480
P478
S654
E705
c
Lobe 1
ξ1
ξ2
Lobe 2
d
Apo
Crystallographic
Inverted
30
25
20
15
10
5
5
10
15
20
25
ξ2 (Å)
ξ1 (Å)

Fig. 4-3

Fig. 4-4: Activation and deactivation of receptors with mutations in Pathway 1. (a) Sites of mutations that were tested functionally. The green-colored sidechains correspond to the WT residues. The tan-colored residues contact bound glutamate directly but were not mutated in the functional tests. (b) Blue circle indicates LBD residues proximal to metastable site 1 (**Fig. 4-2**b) in Lobe 1. Mutants tested include single charge swaps R453D and K458D, and the triple mutant R453A D456A K458A (RDK-AAA). (c) Activation of Lobe 1 mutants (R453D, blue; K458D, green; RDK-AAA, red) by a long pulse of 10 mM glutamate. Solution exchange measured after the experiment is shown as the upper black trace. A typical WT GluA2 response is plotted with a dotted line. The individual 10-90% rise times of the currents (t_{rise}) are shown in the bar chart in the right panel, with the WT mean value as a dashed gray line. Asterisk indicates $P < 0.005$, Student's t-test. (d) Left panel shows deactivation of Lobe 1 mutants in response to $\sim$1 ms pulse of 10 mM glutamate. Color coding is as in (c), with monoexponential fits indicated by open circles. Right panel shows bar chart of individual deactivation decay values. Asterisk indicates $P < 0.005$, Student's t-test. See also **Fig. 4-8**, **Fig. 4-12**, **Fig. 4-13**, and **Table 4.3**.

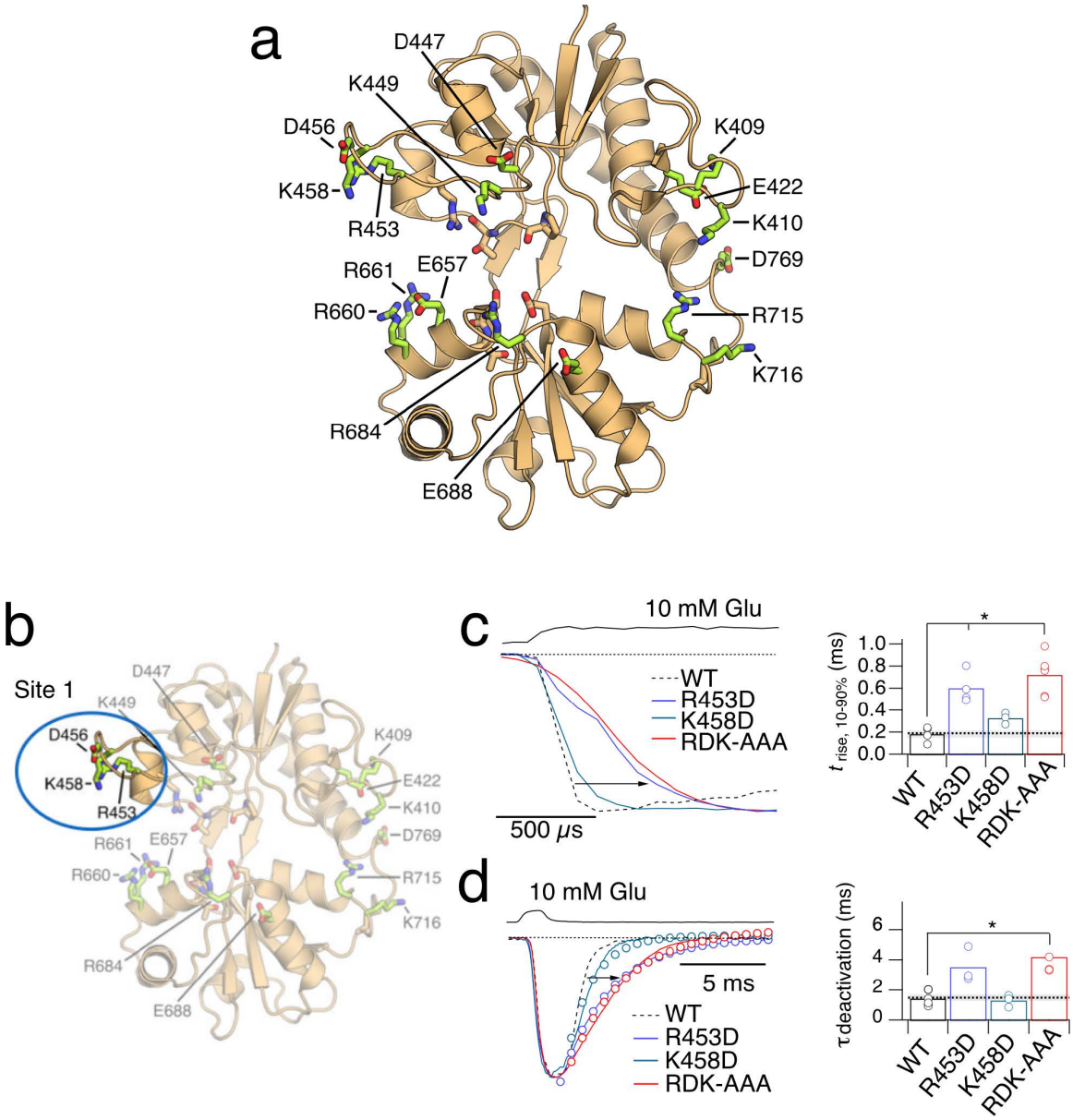
a
D447
K449
D456
K458
R453
K409
E422
K410
D769
R661
E657
R660
R715
K716
R684
E688
b
Site 1
c
10 mM Glu
WT
R453D
K458D
RDK-AAA
500 μs
d
10 mM Glu
5 ms
WT
R453D
K458D
RDK-AAA
t rise, 10-90% (ms)
τ deactivation (ms)
WT
R453D
K458D
RDK-AAA

Fig. 4-4

Fig. 4-5 Activation and deactivation of receptors with mutations in Pathways 1 and 2. (a) Residues in Lobe 2 that interact with the ligand at metastable sites 1 and 2 (**Fig. 4-2**b) included R660E, R661E and E657A R660A R661A (ERR-AAA). (b) Left panel shows activation of Lobe 2 mutants (R660E, green; R661, blue; ERR-AAA, red) in response to a long pulse of 10 mM glutamate. The individual 10-90% rise times of the currents (t_{rise}) are shown in the bar chart in the right panel, with the WT mean value as a dashed gray line, with asterisk indicating $P < 0.05$ vs. WT from t-test. (c) Left panel shows deactivation of Lobe 2 mutants following a 1 ms pulse of 10 mM glutamate, with color coding as in (b). Individual rise times are plotted in the right panel. (d) The affinity for glutamate is unchanged for the ERR-AAA mutant relative to WT. Dose-response curves in glutamate, measured at the peak current response, for WT GluA2 (EC_{50} = 330 ± 90 μM; black circles), and for the mutant ERR-AAA (EC_{50} = 410 ± 30 μM, red circles). By comparing the fits to responses from individual cells, the glutamate EC_{50} for ERR-AAA was indistinguishable from that of WT GluA2 ($p = 0.5$; t-test, $n = 3$). (e) The PMF for the ERR-AAA mutant, contoured at 2.62 kcal/mol, shows a loss of ligand density along Pathways 1 and 2, proximal to helix F. The ERR-AAA ligand-binding pathway, indicated by the red arrow, resembles Pathway 3 of the WT LBD. Interactions between R684 and R675 on Lobe 2 at site 4 are preserved in both the mutant and WT protein. These residues metastably interact with the ligand prior to binding in T_{mut1}. See also **Fig. 4-13**, **Table 4.3**, **Table 4.4**, and Movie S4.

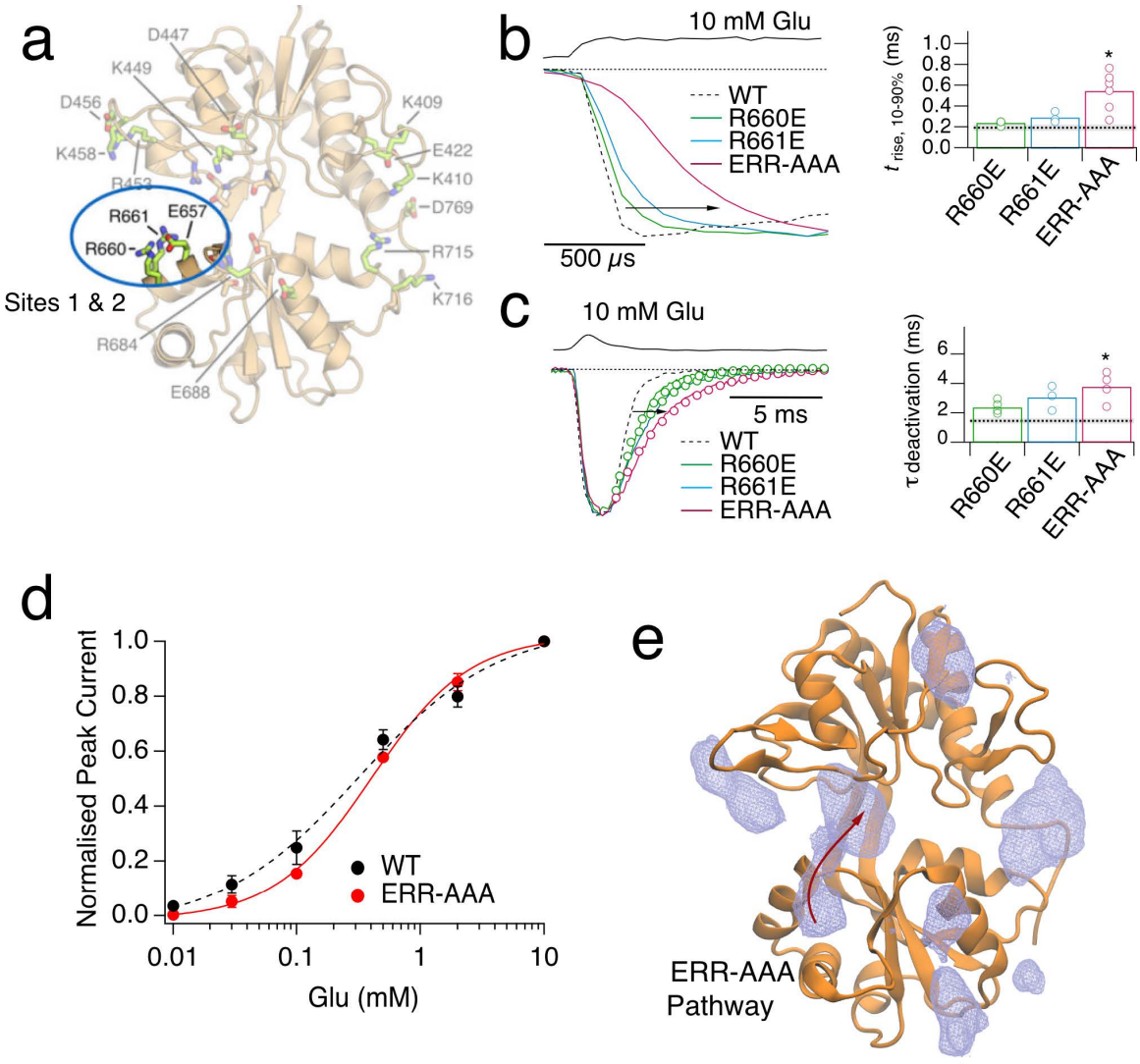
a
D447
K449
D456
K458
R453
R661
E657
R660
Sites 1 & 2
R684
E688
K409
E422
K410
D769
R715
K716
b
10 mM Glu
WT
R660E
R661E
ERR-AAA
500 μs
$t_{rise,\ 10\text{-}90\%}$ (ms)
1.0
0.8
0.6
0.4
0.2
0.0
*
R660E
R661E
ERR-AAA
c
10 mM Glu
5 ms
WT
R660E
R661E
ERR-AAA
τ deactivation (ms)
6
4
2
0
*
R660E
R661E
ERR-AAA
d
Normalised Peak Current
1.0
0.8
0.6
0.4
0.2
0.0
WT
ERR-AAA
0.01
0.1
1
10
Glu (mM)
e
ERR-AAA
Pathway

Fig. 4-5

Fig. 4-6: Activation and deactivation of receptors with mutations in pathway 3 and off-pathway mutants. (a) Blue circles indicate residues in Lobe 1 of the LBD that participate in transferring the ligand from metastable site 3 to site 0 (**Fig. 4-2**b; D447, K449) and that were mutated to alanine (DK-AA). In Lobe 2, two residues that participate in site 3 (R684, E688) were separately mutated to alanine (RE-AA). (b) The activation of receptors in response to a long pulse of 10 mM glutamate were slower than WT GluA2 (dashed line) for both DK-AA (blue trace) and RE-AA (red trace). The upper black trace shows solution exchange. Individual 10-90% rise times are plotted in the right panel with the mean value for WT GluA2 indicated by a dashed line and standard error shaded in light gray. Asterisks indicate $P < 0.01$ vs. WT GluA2, *t*-test, n = 3-6. (c) Deactivation of DK-AA and RE-AA mutants in response to a 1 ms pulse of 10 mM glutamate with color-coding as in (b). Monoexponential fits are represented by open circles. Individual deactivation decay constants are plotted in the bar graph (right panel). The asterisk indicates $P < 0.05$ vs. WT GluA2, *t*-test, n = 3-6. (d) Blue circles indicate positions in the LBD of two triple-alanine mutants, located away from the binding pathways (**Fig. 4-2**a). One set of off- pathway mutants was located in Lobe 1 (K409A, K410A, E422A; KKE-AAA) and one set in Lobe 2 (R715A, K716A, D769A; RKD-AAA), (**Fig. 4-4**a). (e) Left panel shows activation of receptors by long pulses of 10 mM glutamate. Responses for the KKE-AAA (blue trace) and RKD-AAA (green trace) mutants are overlaid with a typical WT GluA2 response as in (c), with individual rise times plotted in the right panel. Asterisk indicates $P < 0.05$. (f) Left panel shows deactivation of off-pathway mutants following a 1 ms pulse of 10 mM glutamate. Color coding as in (e). Individual rise times are plotted in the right panel. Asterisk indicates $P < 0.01$. See also **Fig. 4-14** and **Table 4.3**.

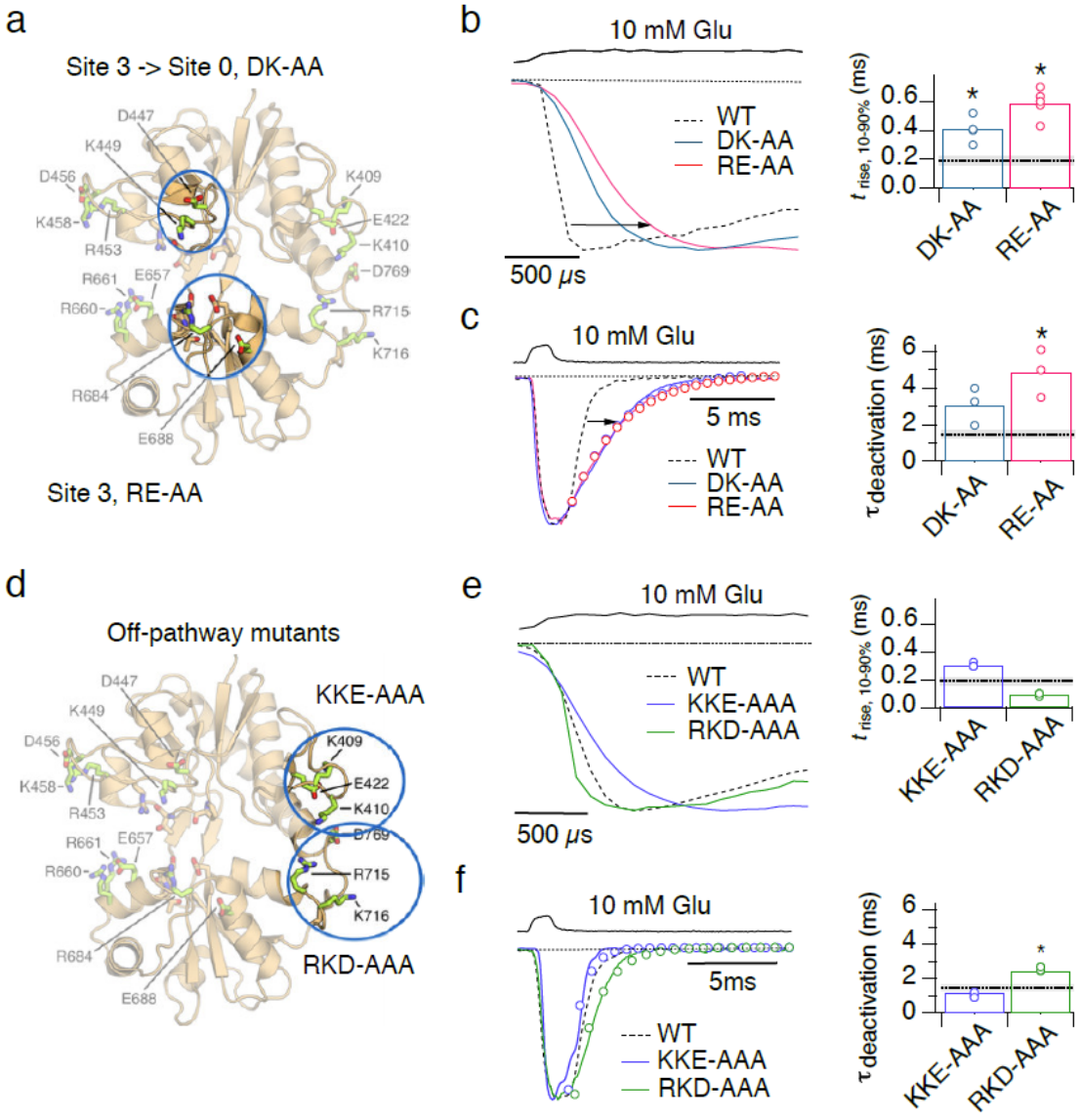
a
Site 3 -> Site 0, DK-AA
D447
K449
D456
K458
R453
K409
E422
K410
D769
R661
E657
R660
R715
K716
R684
E688
Site 3, RE-AA
b
10 mM Glu
WT
DK-AA
RE-AA
500 μs
t rise, 10-90% (ms)
0.6
0.4
0.2
0.0
DK-AA
RE-AA
c
10 mM Glu
5 ms
WT
DK-AA
RE-AA
τ deactivation (ms)
6
4
2
0
DK-AA
RE-AA
d
Off-pathway mutants
KKE-AAA
RKD-AAA
e
10 mM Glu
WT
KKE-AAA
RKD-AAA
500 μs
t rise, 10-90% (ms)
KKE-AAA
RKD-AAA
f
10 mM Glu
5ms
WT
KKE-AAA
RKD-AAA
τ deactivation (ms)
KKE-AAA
RKD-AAA

Fig. 4-6

Fig. 4-7: Kinetic modeling of AMPA receptor currents activated by glutamate recapitulates the effect of slowed binding reactions on receptor activation and deactivation. (a) Kinetic model constructed according to principles outlined in previous studies of GluA2 kinetics [29]. Four glutamate binding sites, two open states (green, *) and five desensitized states (red) are included. Low conductance open states and connections between desensitized states were omitted for simplicity. For each simulation, the association and dissociation rates were multiplied by a common factor, f, to represent the effects of mutants to slow binding rates through disruption of glutamate binding pathways. The rate constants were as follows: beta = 5000 s^{-1}, alpha = 3000 s^{-1}, $k+ = 5 \times 10^6$ M^{-1} s^{-1}, k- = 10000 s^{-1} , $d+$ = 250 s^{-1}, d- = 60 s^{-1} , $d_0+ = 1$ s^{-1} , d_0- = 9 s^{-1} . The conductance of the open state A4R* was set at twice that of A3R*. (b) Example simulated currents (normalized to the maximum possible response) for a realistic concentration jump of 800 μs with rise time of 300 μs, using the model in panel a. Six color-coded current profiles generated using the RCJ scripts (see Methods) are shown, with the binding rate factor ranging from 2 to 0.05. Note that deactivation is more strongly affected than activation (because efficacy for channel opening is > 1). Also, for fast binding reactions, the rise time is faster than the solution exchange. (c) Kinetic measurements from GluA2 mutants and wild-type (black circles, WT GluA2 marked) were well described by a linear fit with slope of $\sim$8 with intercept close to the origin. 95% confidence intervals for the line are shown as grey dashed curves. Uncertainties in both abscissa and ordinate were used for the fit (ODR 2 in Igor 7). Off-pathway control mutants (red circles) lie away from this curve. Data from simulations with different rise times for the glutamate pulse are plotted as open symbols, with the relevant f value for each group of simulations indicated with a dotted line. The colors of these symbols relates to the simulated 10-90% solution exchange time seen by receptors. All solution exchange rates predict the similar steep, approximately linear relations between activation time (10-90%) and decay time constant. The simulated kinetics span a similar range to the electrophysiological recordings of pathway-disruption mutants. The best agreement between simulation and experiment comes from solution exchange times

at an intact patch in the physically plausible range of 200-400 μs.

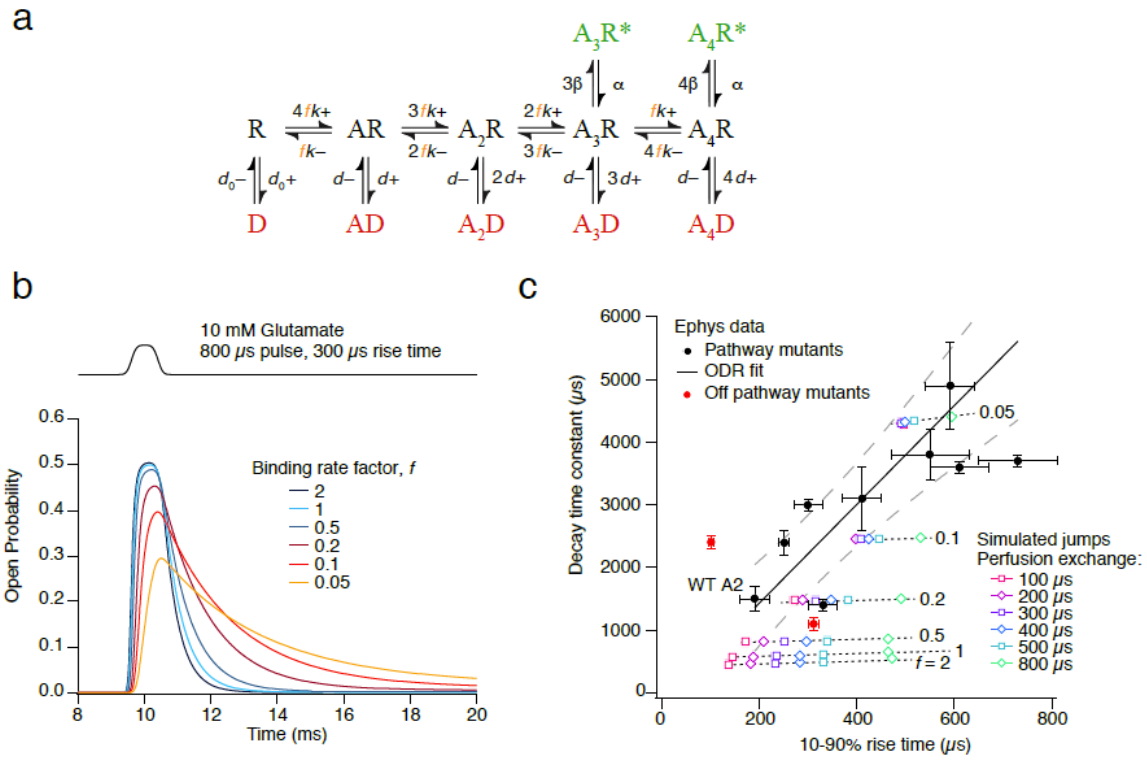
a
A3R*
A4R*
3β
α
4β
α
R
4fk+
fk–
AR
3fk+
2fk–
A2R
2fk+
3fk–
A3R
fk+
4fk–
A4R
d0–
d0+
d–
d+
d–
2d+
d–
3d+
d–
4d+
D
AD
A2D
A3D
A4D
b
10 mM Glutamate
800 μs pulse, 300 μs rise time
Open Probability
Binding rate factor, f
2
1
0.5
0.2
0.1
0.05
Time (ms)
c
Ephys data
Pathway mutants
ODR fit
Off pathway mutants
Decay time constant (μs)
0.05
0.1
0.2
0.5
1
f = 2
WT A2
Simulated jumps
Perfusion exchange:
100 μs
200 μs
300 μs
400 μs
500 μs
800 μs
10-90% rise time (μs)

Fig. 4-7

Fig. 4-8 Sites of metastable protein-ligand interactions. (a) Sequence alignment of iGluR LBDs across AMPA, kainate, and NMDA receptors in rat. Binding pocket residues are boxed in red. Residues that make metastable interactions are boxed in yellow. Off-pathway residues that were mutated are boxed in black. (b) Sites of mutation in a tetrameric GluA2 receptor. Each sphere corresponds to a site of mutation shown in **Fig. 4-4**a in a tetrameric receptor. R453, D456, and K458 are in red; E657, R660, and R661 are in magenta; D447 and K449 are in yellow; R684 and E688 are in green; K409, K410, and E422 are in blue; R715, K716, and D769 are in violet. The mutations occur in all subunits of the receptor, but they are shown in only one subunit for clarity. The structure shown is PDB ID 4U2P [49].

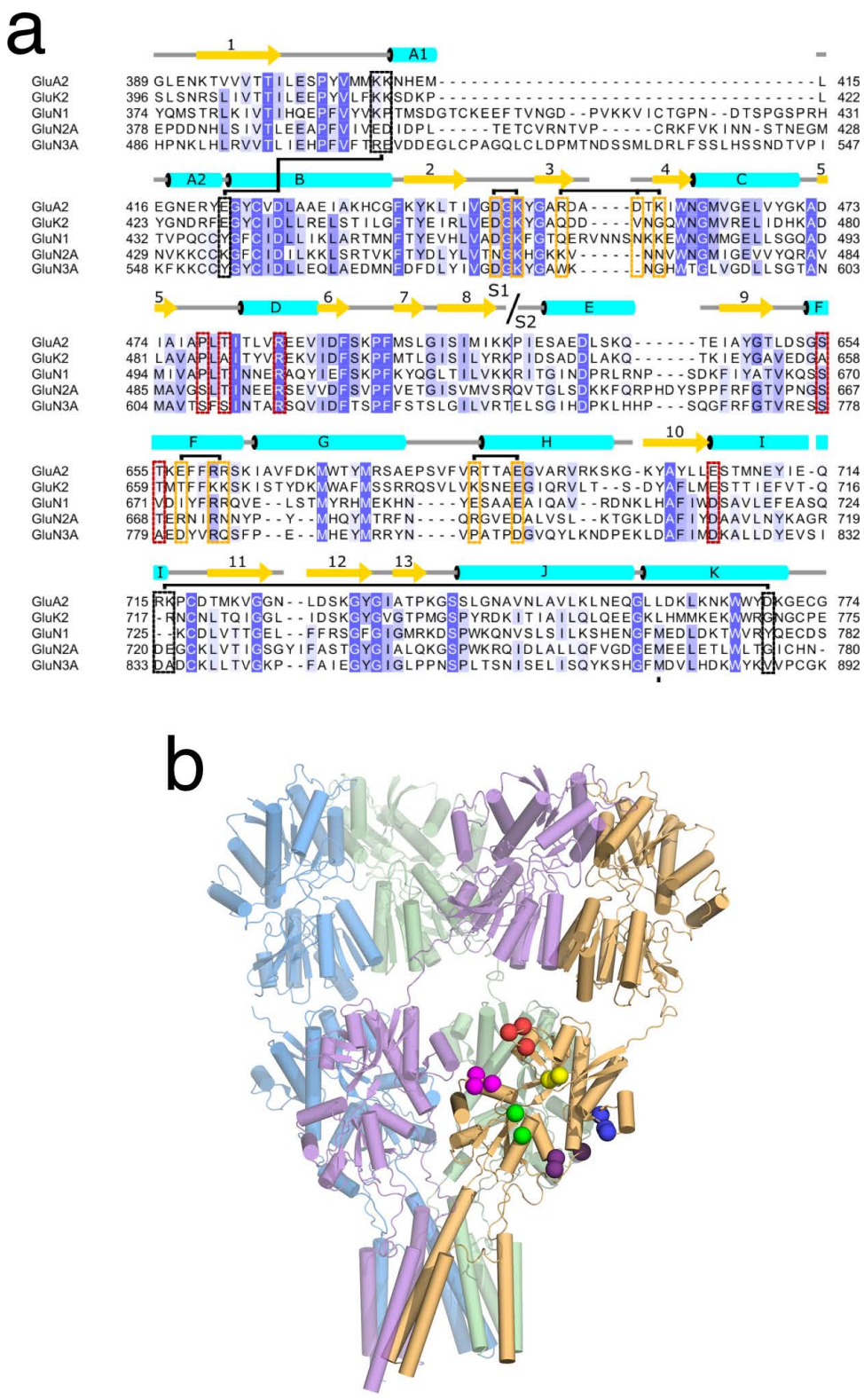

Fig. 4-8

Fig. 4-9 Error analysis of the ligand density PMF. Related to Figure 2. (a-f) The 3D PMF (blue) is shown at contour levels ranging from 0.6 kcal/mol to 3.6 kcal/mol in increments of 0.6 kcal/mol. (g-l) The statistical uncertainty in the 3D PMF determined using the approach of block averaging. 10 blocks were used. Contours of the 3D standard deviation (red) are shown from 0.2 kcal/mol to 1.2 kcal/mol in increments of 0.2 kcal/mol. Protein-ligand interactions (a-d) have uncertainties ranging from approximately ± 0.8 to ± 1.2 kcal/mol (g-l), whereas sites in bulk solvent (e, f) have lower uncertainties of approximately ± 0.2 to ± 0.6 kcal/mol (j-l).

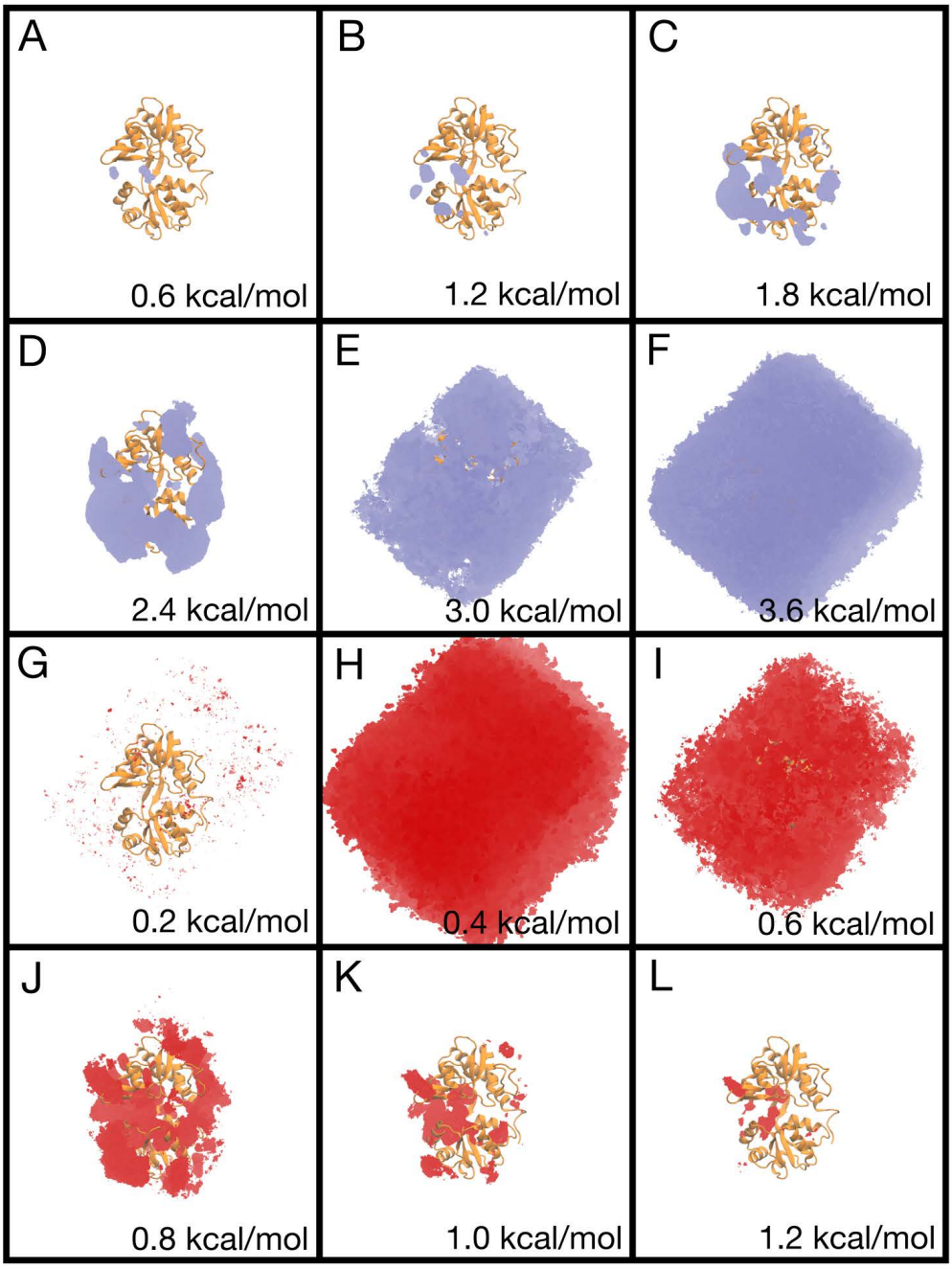
A
0.6 kcal/mol
B
1.2 kcal/mol
C
1.8 kcal/mol
D
2.4 kcal/mol
E
3.0 kcal/mol
F
3.6 kcal/mol
G
0.2 kcal/mol
H
0.4 kcal/mol
I
0.6 kcal/mol
J
0.8 kcal/mol
K
1.0 kcal/mol
L
1.2 kcal/mol

Fig. 4-9

Fig. 4-10: Glutamate binding in the inverted conformation. Time points (lower left of each panel) are relative to the start of Movie S2. (a) Prior to ligand binding, the LBD is closed; (ξ_1, ξ_2) = (11.0, 11.7 Å). The ligand's γ-carboxylate contacts R660. (b) Close-up view of (a). (c-d) The ligand moves away from R485 as it is passed from R660 to R661. Interactions between the ligand and helix F residues switch from the γ-carboxylate at R660 to the α-carboxylate at R661. (e) The LBD opens, (ξ_1, ξ_2) = (15.0, 11.2 Å), to allow ligand entry into the binding pocket. (f-g) The ligand metastably bridges Lobes 1 and 2 as its α-carboxylate contacts R485. (h) Interactions with helix F residues are broken, and the ligand moves into the binding pocket, tethered to R485. (i) Glutamate adopts the inverted bound conformation. (j-k) The ligand's amide contacts E705. A hydrogen bonding network is formed involving P478, a water molecule, and the ligand's amide, stabilizing the inverted conformation. (l) Expanded view of (k). The LBD remains slightly open around the inverted conformation; (ξ_1, ξ_2) = (13.2, 11.1 Å). This trajectory corresponds to T_{mon1} in **Table 4.1**.

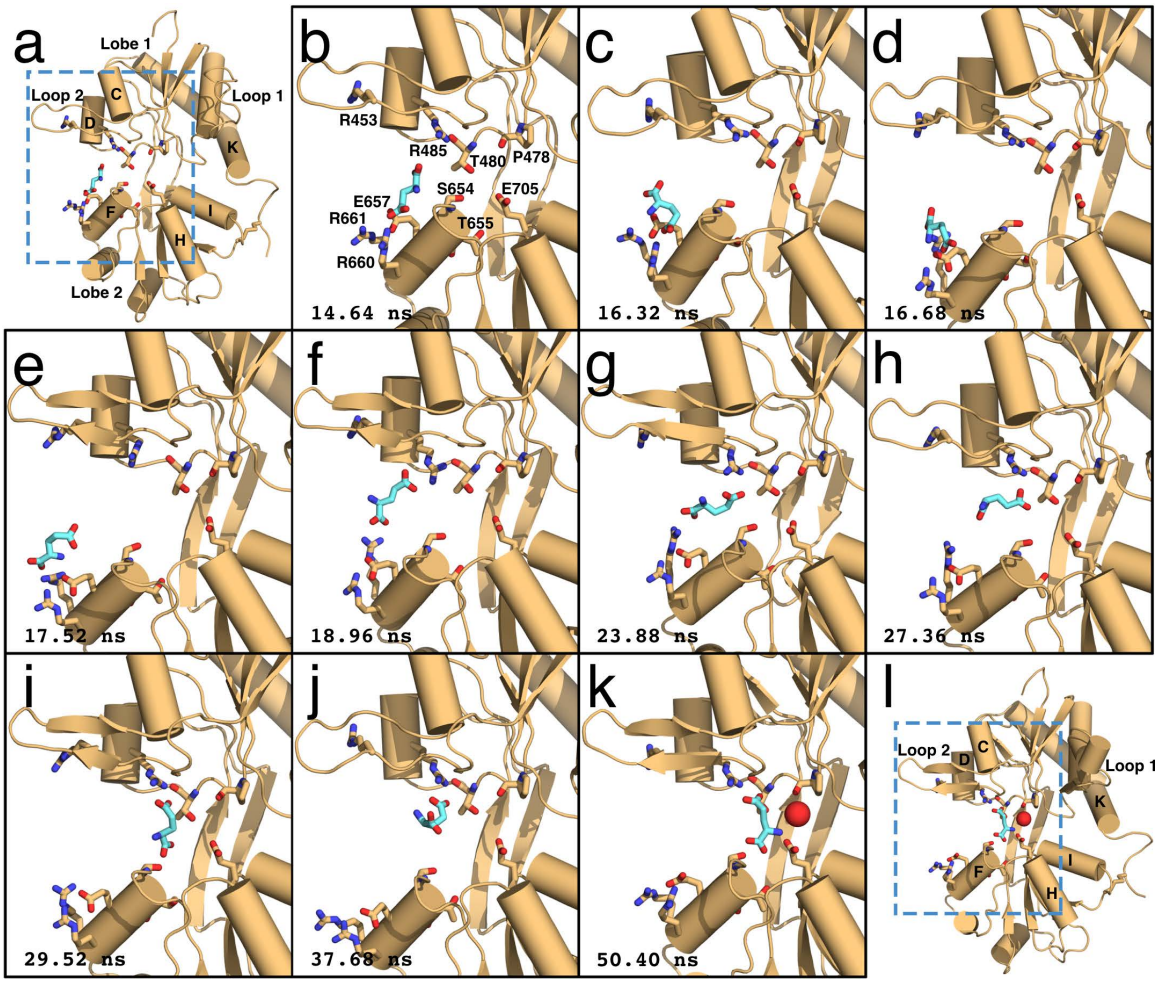
a
Lobe 1
Loop 2
C
Loop 1
D
K
F
I
H
Lobe 2
b
R453
R485
T480
P478
E657
S654
E705
R661
T655
R660
14.64 ns
c
16.32 ns
d
16.68 ns
e
17.52 ns
f
18.96 ns
g
23.88 ns
h
27.36 ns
i
29.52 ns
j
37.68 ns
k
50.40 ns
l
Loop 2
D
C
Loop 1
K
F
I
H

Fig. 4-10

Fig. 4-11: Interconversion between bound ligand conformations. Time points (lower left of each panel) are relative to the start of Movie S3. (a) Initially, the ligand binds in the crystallographic conformation. (b) Close-up view of (a); the ligand's α-carboxylate contacts R485, whereas the amide is coordinated by P478 and E705. (c-d) The γ-carboxylate swings out of the binding pocket to contact R453. Interactions between the ligand's amide with P478 and E705 are broken. (e-f) The γ-carboxylate rotates freely in the binding cleft while the α-carboxylate remains tethered to R485. (g) Contacts between the α-carboxylate and R485 are severed as the ligand repositions. (h) The γ-carboxylate of the ligand contacts R485. (i) The ligand swings into the binding pocket. (j) Glutamate adopts the inverted bound conformation. (k) The ligand's amide contacts E705, and P478 forms a hydrogen bond with a water molecule that stabilizes the inverted conformation. (l) Expanded view of (k). This trajectory corresponds to T_{dim2} in Table S1.

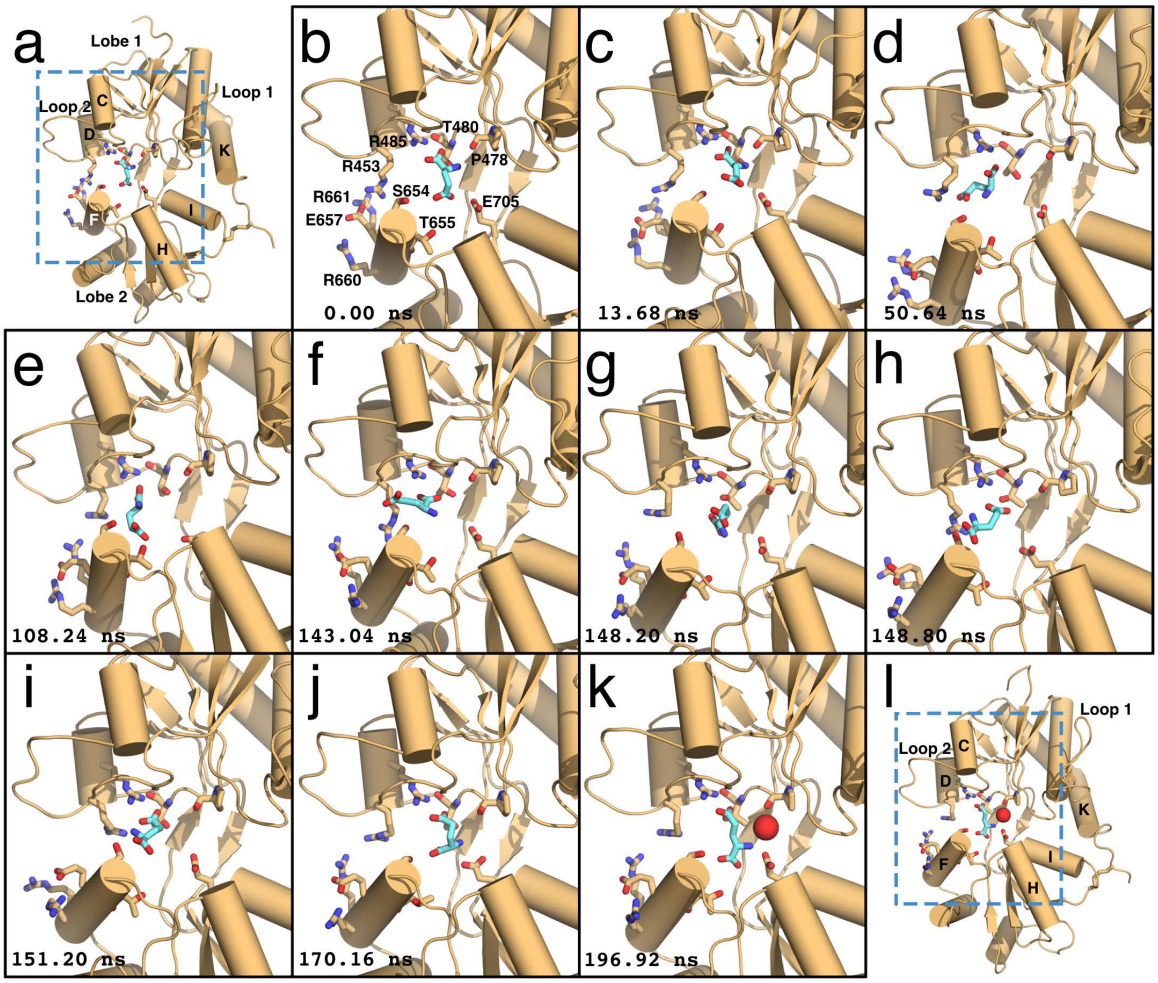
a
Lobe 1
Loop 1
Loop 2
C
D
K
F
I
H
Lobe 2
b
T480
R485
P478
R453
S654
R661
E705
E657
T655
R660
0.00 ns
c
13.68 ns
d
50.64 ns
e
108.24 ns
f
143.04 ns
g
148.20 ns
h
148.80 ns
i
151.20 ns
j
170.16 ns
k
196.92 ns
l
Loop 1
Loop 2
C
D
K
F
I
H

Fig. 4-11

Fig. 4-12: Slower activation and deactivation are retained in high glutamate. (a) Normalized current responses corresponding to activation by a long pulse of 50 mM glutamate are shown for the RDK-AAA mutant (red) with WT GluA2 (dashed black trace). The upper trace shows the application of glutamate. Rise times (10-90%) from individual patches are shown in the bar graph (right panel). ($p = 0.01$ vs. WT; t-test, $n = 3$). (b) Monoexponential decays for WT GluA2 (dashed black line) and RDK-AAA (red line) were fitted for the deactivation in response to a 1 ms pulse of 50 mM glutamate (open circles). The upper trace shows the open tip response. Deactivation time constants from individual patches are plotted in the bar graph (right panel; $p = 0.0003$; t-test, $n = 3$). (c) Desensitization time constants were not altered by the RDK-AAA mutant. Monoexponential fits as in (b). Desensitization time constants from individual patches plotted in the bar graph (right panel) revealed no difference between WT and RDK-AAA ($p = 0.3$; t-test, $n = 3$).

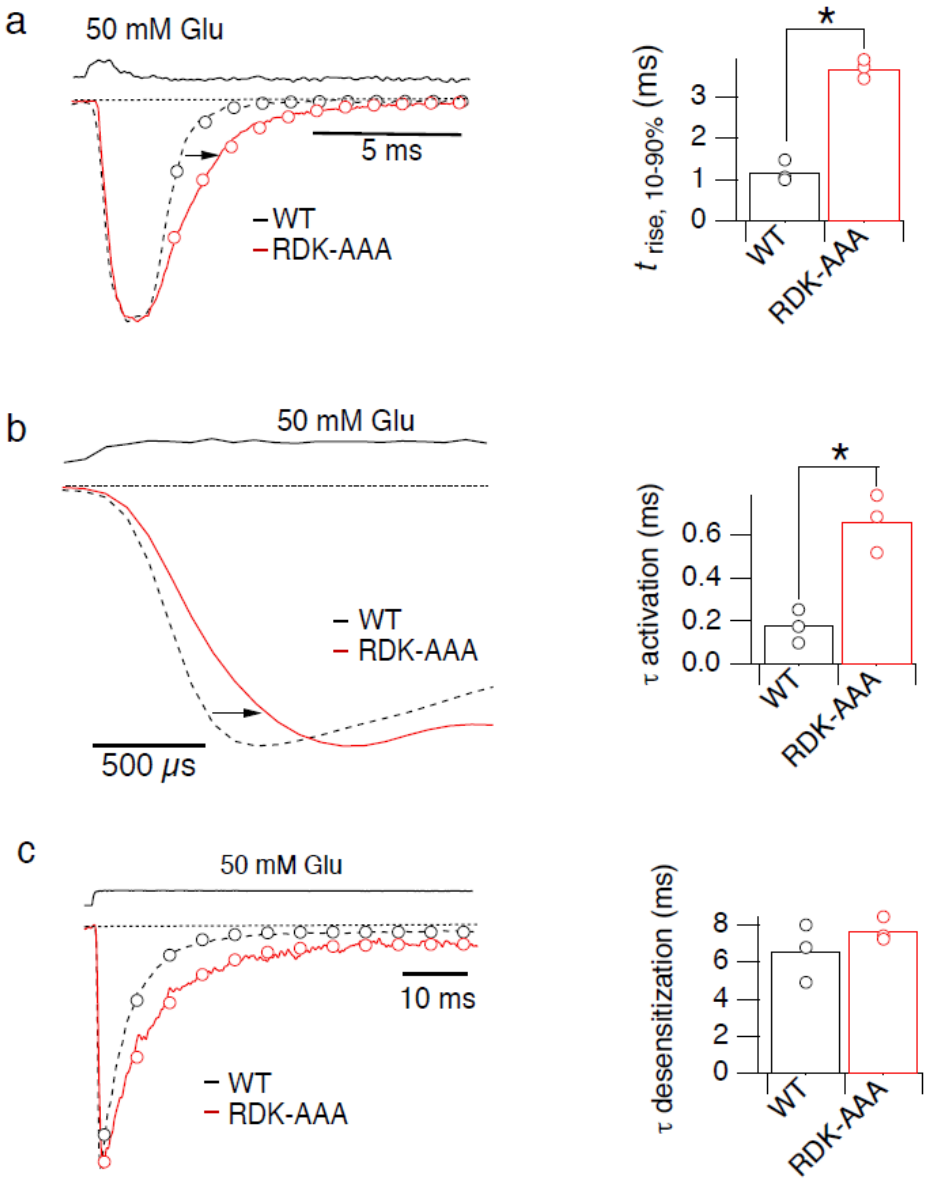
a
50 mM Glu
5 ms
WT
RDK-AAA
*
$t_{rise,\ 10\text{-}90\%}$ (ms)
3
2
1
0
WT
RDK-AAA
b
50 mM Glu
WT
RDK-AAA
500 μs
*
τ activation (ms)
0.6
0.4
0.2
0.0
WT
RDK-AAA
c
50 mM Glu
10 ms
WT
RDK-AAA
τ desensitization (ms)
8
6
4
2
0
WT
RDK-AAA

Fig. 4-12

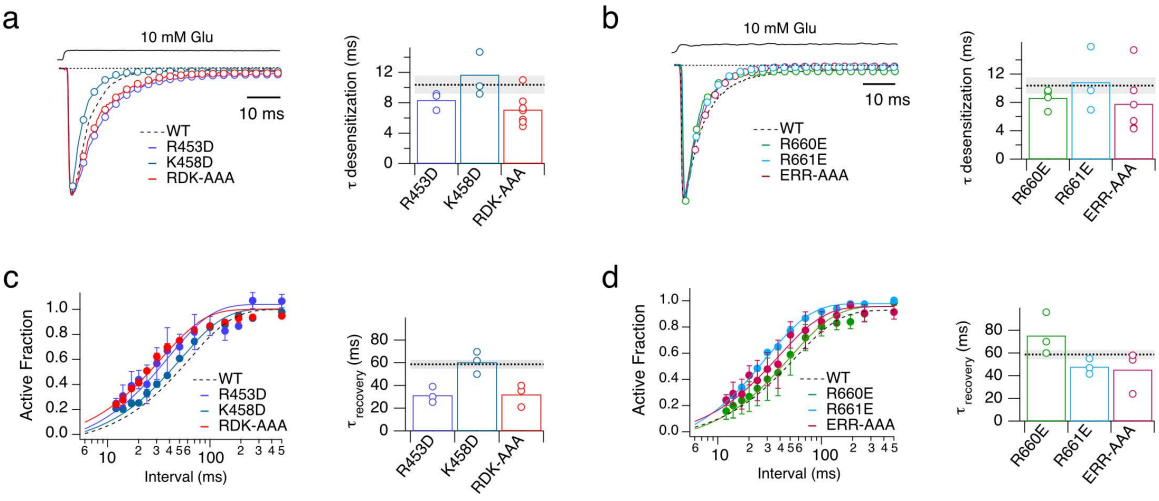

Fig. 4-13: **Desensitization and recovery from desensitization of receptors mutated at metastable sites 1 and 2.** (a) Monoexponential fits to desensitization decays of R453D, K458D and the R453D, D456A, K458A (RDK-AAA) mutants in response to 10 mM glutamate were similar to WT ($p > 0.03$; t-test, $n = 3$ - 7). A typical WT GluA2 response is shown as a dashed line. Decay constants for individual patches are plotted in the right panel, with the average value for WT GluA2 indicated by a black dashed line, with standard error shaded in light grey. (b) As for (a), but for the R660E, R661E and E657A, R660A, R661A triple mutants. Fits were similar to that of WT ($p > 0.2$; t-test, $n = 3$). (c) Left panels show fits to pooled responses ("Active Fraction") at increasing intervals after a long pulse of 10 mM glutamate for WT and the different mutants. The R453D and RDK-AAA mutants recovered from desensitization about twice as fast as WT ($p < 0.02$; t-test, $n = 3$). Recovery time constants for individual patches, with the WT mean value and standard error indicated as in (a-b). (d) Recovery data for the R660E, R661E and ERR-AAA mutants presented as in (c). Recovery time constants for these mutants were not significantly different from that of WT ($p > 0.1$; t-test, $n = 3$).

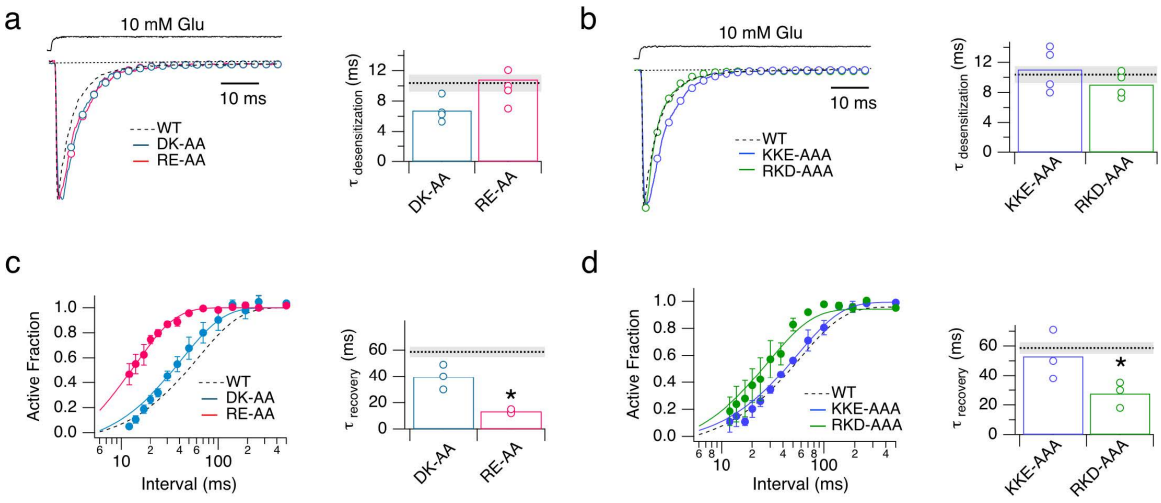

Fig. 4-14: **Desensitization and recovery from desensitization of receptors with mutations in pathway 3 and in off-pathway regions.** (a) Monoexponential fits to desensitization decays of DK-AA (D447A K449A) and RE-AA (R684A, E688A) mutants in response to 10 mM glutamate were similar to WT (p = 0.03 for DK-AA, p = 0.8 for RE-AA; *t*-test, n = 3 - 7). Fits are represented by open circles. A typical WT GluA2 response is overlaid (black dashed line). Decay constants for individual patches are plotted in the right panel, with the average value for WT GluA2 indicated by a black dotted line, with its standard error shaded in light grey. (b) As for (a), but for the KKE-AAA and RKD-AAA control mutants (P = 0.7 and 0.4 vs. WT, respectively; *t*-test). (c) Left panels show fits to pooled responses ("Active Fraction") at increasing intervals after a long pulse of 10 mM glutamate for WT and the DK-AA and RE-AA mutants. The RE-AA mutant gave very small currents recovered from desensitization about four times as fast as WT (p = 0.008; *t*-test, n = 3 - 5). Recovery time constants for individual patches, with the WT mean value and standard error indicated as in (a-b). (d) Recovery data for the KKE-AAA and RKD-AAA mutants (P = 0.6 and 0.01 vs. WT, respectively; *t*-test), presented as in (c).

	No. of ligands	Event [a]	Apo, prior [b]	ligand docked [c]	ligand conformation	(ξ1, ξ2) median[d]	Apo, post [e]	Previous Trajectory
T_{dim1}	20	A	0.93	2.49	crystallographic	11.7, 11.1	N/A	N/A
T_{dim2}	20	A	1.39	0.88	crystallographic → inverted	13.1, 11.5	N/A	N/A
T_{dim3}	20	A	N/A	3.12	inverted	12.9, 11.6	N/A	T_{dim6}
T_{dim4}	20	A	0.08	1.4	inverted	12.7, 11.2	N/A	T_{dim5}
T_{dim5}	20	D	N/A	2.49	crystallographic	11.7, 11.1	0.08	T_{dim1}
T_{dim6}	20	D	N/A	0.88	inverted	13.1, 11.5	N/A	T_{dim2}
T_{dim7}	20	D	N/A	1.4	inverted	12.7, 11.2	0.48	T_{dim4}
T_{dim8}	2	D	N/A	1.44	crystallographic	11.5, 11.1	4.98	N/A
T_{dim9}	2	D	N/A	3.57	crystallographic	11.7, 10.8	2.85	N/A
T_{mon1}	1	A	2.94	2.04	inverted	12.6, 11.1	N/A	T_{mon4}
T_{mon2}	10	A	1.25	0.78	inverted	14.1, 11.7	N/A	N/A
T_{mon3}	10	D	N/A	0.78	inverted	14.1, 11.7	0.88	T_{mon2}
T_{mon4}	1	D	N/A	1.25	crystallographic	11.1, 11.4	2.94	N/A
T_{mon5}	1	D	N/A	0.98	crystallographic	11.2, 10.7	9.32	N/A
T_{mon6}	1	D	N/A	5.52	crystallographic → inverted	12.4, 11.7	5.21	N/A
T_{mon7}	1	D	N/A	1.84	crystallographic	11.5, 11.0	7.2	N/A
T_{mon8}	1	D	N/A	1.85	crystallographic	11.8, 11.2	6.51	N/A

Table 4.1: **Ligand-binding trajectories for the dimer and monomer systems.**

[a] Binding events are labeled A for association and D for dissociation

[b] μs in the apo conformation prior to the binding event

[c] μs in the docked ligand conformation

[d] units are in Ångstroms

[e] μs in the apo conformation after the binding event

No. of subunits	No. of Ligands	Free Ligand Concentration (mM)	Time spent in bulk solvent (μs)	Binding Events
1	1	3.9	31.18	1
2	1	3.5	2.13	0
2	2	6.9	2.85	0
1	9	35.1	0.78	0
1	10	38.9	2.13	1
2	18	63.5	3.43	2
2	19	67.1	1.02	1
2	20	70.6	0.93	1

Table 4.2: **Quantities used in the calculation of $k_{\mathbf{on}} = (1.4 \times 10^7\ M^{-1}s^{-1}$.** $k_{\text{on}} = \frac{N_b}{\sum_i \frac{t_i[L_i]}{s_i}}$, where N_b is the total number of binding events; t_i is the time the ligand spends in bulk solvent, $[L_i]$ is the free ligand concentration, s_i is the number of protein subunits, and i is summed over all simulation systems.

Mutant / Condition	$t_{10\text{-}90\%\ rise}$, µs (*n*)	*P*	τ_{deact}, ms (*n*)	*P*	τ_{des}, ms (*n*)	*P*	τ_{rec}, ms (*n*)	*P*
WT GluA2	190 ± 30 (4)		1.5 ± 0.2 (6)		10.4 ± 1.1 (8)		59 ± 4 (3)	
R453D	610 ± 60 (4)	0.005	3.6 ± 0.1 (3)	0.08	8.4 ± 0.6 (3)	0.2	31 ± 3 (3)	0.01
K458D	330 ± 30 (3)	0.03	1.4 ± 0.1 (3)	0.7	11.7 ± 1.3 (3)	0.5	60 ± 5 (3)	0.8
R453A, D456A, K458A (RDK-AAA)	730 ± 80 (6)	0.0003	3.7 ± 0.1 (4)	0.003	7.1 ± 0.7 (7)	0.03	32 ± 5 (3)	0.02
R660E	250 ± 10 (4)	0.2	2.4 ± 0.2 (4)	0.02	8.6 ± 0.6 (4)	0.2	75 ± 9 (3)	0.3
R661E	300 ± 30 (3)	0.07	3 ± 0.1 (3)	0.07	10.8 ± 2.3 (3)	0.9	48 ± 3 (3)	0.1
E657A, R660A, R661A (ERR-AAA)	550 ± 80 (6)	0.004	3.8 ± 0.4 (4)	0.01	7.8 ± 1.7 (6)	0.2	45 ± 9 (3)	0.3
D447A, K449A (DK-AA)	410 ± 40 (4)	0.01	3.1 ± 0.5 (3)	0.1	6.8 ± 0.8 (4)	0.03	40 ± 5 (3)	0.06
R684A, E688A (RE-AA)	590 ± 50 (5)	0.0002	4.9 ± 0.7 (3)	0.04	10.8 ± 1.3 (5)	0.8	13 ± 1 (3)	0.008
K409A, K410A, E422A (KKE-AAA)	310 ± 10 (3)	0.04	1.1 ± 0.1 (4)	0.1	11.1 ± 1.3 (4)	0.7	53 ± 8 (3)	0.6
R715A K716A D769A (RKD-AAA)	100 ± 5 (4)	0.07	2.4 ± 0.1 (4)	0.005	9 ± 0.7 (4)	0.4	28 ± 4 (3)	0.01
WT GluA2 (50 mM Glu)	180 ± 40 (3)		1.2 ± 0.1 (3)		6.6 ± 0.8 (3)			
RDK-AAA (50 mM Glu)	660 ± 70 (3)	0.01	3.7 ± 0.1 (3)	0.0003	7.7 ± 0.3 (3)	0.3		

Table 4.3: **Kinetic properties of WT GluA2 and mutants.** All experiments were done with 10 mM glutamate except where noted. The number of experiments is indicated in brackets. $t_{10-90\%\text{rise}}$, rise time; τ_{deact}, time constant from single exponential fits to the deactivation decay; τ_{des}, time constant from single exponential fits to the desensitization decay; τ_{rec}, recovery time constant from Hodgkin-Huxley fits with $h = 2$ (see Materials and Methods). P values are from two-tailed Student's t-test against WT values.

	No. of ligands	Event [a]	Apo, prior [b]	ligand docked [c]	ligand conformation	(ξ1, ξ2) median[d]	Apo, post [e]	Previous Trajectory
T_{mut1}	10	A	0.06	0.31	inverted	15.3, 11.8	N/A	N/A
T_{mut2}	10	A	0.10	0.75	inverted	12.8, 11.5	N/A	T_{mut3}
T_{mut3}	10	D	N/A	0.31	inverted	15.3, 11.8	0.10	T_{mut1}
T_{mut4}	10	D	N/A	0.75	inverted	12.8, 11.5	N/A	N/A

Table 4.4: **Ligand-binding trajectories for the ERR-AAA mutant LBD.**

[a] Binding events are labeled A for association and D for dissociation

[b] μs in the apo conformation prior to the binding event

[c] μs in the docked ligand conformation

[d] units are in Ångstroms

[e] μs in the apo conformation after the binding event

References

[1] Attwell, D. and Gibb, A. Neuroenergetics and the kinetic design of excitatory synapses. *Nature Reviews. Neuroscience*, 6(11):841–849, November 2005.

[2] Lisman, J. E., Raghavachari, S., and Tsien, R. W. The sequence of events that underlie quantal transmission at central glutamatergic synapses. *Nature Reviews. Neuroscience*, 8(8):597–609, August 2007.

[3] Baranovic, J. and Plested, A. J. R. How to build the fastest receptor on earth. *Biological Chemistry*, 397(3):195–205, March 2016.

[4] Traynelis, S. F., Wollmuth, L. P., McBain, C. J., Menniti, F. S., Vance, K. M., Ogden, K. K., Hansen, K. B., Yuan, H., Myers, S. J., and Dingledine, R. Glutamate receptor ion channels: structure, regulation, and function. *Pharmacological Reviews*, 62(3):405–496, September 2010.

[5] Mayer, M. L. Emerging models of glutamate receptor ion channel structure and function. *Structure (London, England: 1993)*, 19(10):1370–1380, October 2011.

[6] Mayer, M. L. Structural biology of glutamate receptor ion channel complexes. *Current Opinion in Structural Biology*, 41:119–127, December 2016.

[7] Arinaminpathy, Y., Sansom, M. S. P., and Biggin, P. C. Molecular dynamics simulations of the ligand-binding domain of the ionotropic glutamate receptor GluR2. *Biophysical Journal*, 82(2):676–683, February 2002.

[8] Speranskiy, K. and Kurnikova, M. On the Binding Determinants of the Glutamate Agonist with the Glutamate Receptor Ligand Binding Domain. *Biochemistry*, 44(34):11508–11517, August 2005.

[9] Mendieta, J., Gago, F., and Ramirez, G. Binding of 5'-GMP to the GluR2 AMPA receptor: insight from targeted molecular dynamics simulations. *Biochemistry*, 44(44):14470–14476, November 2005.

[10] Lau, A. Y. and Roux, B. The Free Energy Landscapes Governing Conformational Changes in a Glutamate Receptor Ligand-Binding Domain. *Structure (London, England : 1993)*, 15(10):1203–1214, October 2007.

[11] Mamonova, T., Yonkunas, M. J., and Kurnikova, M. G. Energetics of the cleft closing transition and the role of electrostatic interactions in conformational rearrangements of the glutamate receptor ligand binding domain. *Biochemistry*, 47(42):11077–11085, October 2008.

[12] Postila, P. A., Swanson, G. T., and Pentikainen, O. T. Exploring kainate receptor pharmacology using molecular dynamics simulations. *Neuropharmacology*, 58(2):515–527, February 2010.

[13] Sahai, M. A. and Biggin, P. C. Quantifying water-mediated protein-ligand interactions in a glutamate receptor: a DFT study. *The Journal of Physical Chemistry. B*, 115(21):7085–7096, June 2011.

[14] Lau, A. Y. and Roux, B. The hidden energetics of ligand binding and activation in a glutamate receptor. *Nature Structural & Molecular Biology*, 18(3):283–287, March 2011.

[15] Okada, O., Odai, K., Sugimoto, T., and Ito, E. Molecular dynamics simulations for glutamate-binding and cleft-closing processes of the ligand-binding domain of GluR2. *Biophysical Chemistry*, 162:35–44, March 2012.

[16] Wolter, T., Steinbrecher, T., and Elstner, M. Computational Study of Synthetic Agonist Ligands of Ionotropic Glutamate Receptors. *PLOS ONE*, 8(3):e58774, March 2013.

[17] Yao, Y., Belcher, J., Berger, A. J., Mayer, M. L., and Lau, A. Y. Conformational analysis of NMDA receptor GluN1, GluN2, and GluN3 ligand-binding domains reveals subtype-specific characteristics. *Structure (London, England: 1993)*, 21(10):1788–1799, October 2013.

[18] Dai, J. and Zhou, H.-X. Semiclosed Conformations of the Ligand-Binding Domains of NMDA Receptors during Stationary Gating. *Biophysical Journal*, 111(7):1418–1428, October 2016.

[19] Shaw, D. E., Dror, R. O., Salmon, J. K., Grossman, J. P., Mackenzie, K. M., Bank, J. A., Young, C., Deneroff, M. M., Batson, B., Bowers, K. J., Chow, E., Eastwood, M. P., Ierardi, D. J., Klepeis, J. L., Kuskin, J. S., Larson, R. H., Lindorff-Larsen, K., Maragakis, P., Moraes, M. A., Piana, S., Shan, Y., and Towles, B. Millisecond-scale Molecular Dynamics Simulations on Anton. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*, SC '09, pages 39:1–39:11, New York, NY, USA, 2009. ACM.

[20] Cheng, Q., Du, M., Ramanoudjame, G., and Jayaraman, V. Evolution of glutamate interactions during binding to a glutamate receptor. *Nature Chemical Biology*, 1(6):329–332, November 2005.

[21] Armstrong, N. and Gouaux, E. Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron*, 28(1):165–181, October 2000.

[22] Dror, R. O., Pan, A. C., Arlow, D. H., Borhani, D. W., Maragakis, P., Shan, Y., Xu, H., and Shaw, D. E. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proceedings of the National Academy of Sciences of the United States of America*, 108(32):13118–13123, August 2011.

[23] Clements, J. D., Lester, R. A., Tong, G., Jahr, C. E., and Westbrook, G. L. The time course of glutamate in the synaptic cleft. *Science (New York, N.Y.)*, 258(5087):1498–1501, November 1992.

[24] Rudolph, S., Tsai, M.-C., von Gersdorff, H., and Wadiche, J. I. The ubiquitous nature of multivesicular release. *Trends in Neurosciences*, 38(7):428–438, July 2015.

[25] Weston, M. C., Gertler, C., Mayer, M. L., and Rosenmund, C. Interdomain interactions in AMPA and kainate receptors regulate affinity for glutamate. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 26(29):7650–7658, July 2006.

[26] Abele, R., Keinanen, K., and Madden, D. R. Agonist-induced isomerization in a glutamate receptor ligand-binding domain. A kinetic and mutagenetic analysis. *The Journal of Biological Chemistry*, 275(28):21355–21363, July 2000.

[27] Yu, A., Alberstein, R., Thomas, A., Zimmet, A., Grey, R., Mayer, M. L., and Lau, A. Y. Molecular lock regulates binding of glycine to a primitive NMDA receptor. *Proceedings of the National Academy of Sciences*, 113(44):E6786–E6795, November 2016.

[28] Rosenmund, C., Stern-Bach, Y., and Stevens, C. F. The tetrameric structure of a glutamate receptor channel. *Science (New York, N.Y.)*, 280(5369):1596–1599, June 1998.

[29] Robert, A., Armstrong, N., Gouaux, J. E., and Howe, J. R. AMPA Receptor Binding Cleft Mutations That Alter Affinity, Efficacy, and Recovery from Desensitization. *Journal of Neuroscience*, 25(15):3752–3762, April 2005.

[30] Lape, R., Plested, A. J. R., Moroni, M., Colquhoun, D., and Sivilotti, L. G. The α1k276e Startle Disease Mutation Reveals Multiple Intermediate States in the Gating of Glycine Receptors. *Journal of Neuroscience*, 32(4):1336–1352, January 2012.

[31] Sylantyev, S., Savtchenko, L. P., Niu, Y.-P., Ivanov, A. I., Jensen, T. P., Kullmann, D. M., Xiao, M.-Y., and Rusakov, D. A. Electric Fields Due to Synaptic Currents Sharpen Excitatory Transmission. *Science (New York, N.Y.)*, 319(5871):1845–1849, March 2008.

[32] McCammon, J. A. Darwinian biophysics: Electrostatics and evolution in the kinetics of molecular binding. *Proceedings of the National Academy of Sciences*, 106(19):7683–7684, May 2009.

[33] Fiser, A. and Sali, A. ModLoop: automated modeling of loops in protein structures. *Bioinformatics (Oxford, England)*, 19(18):2500–2501, December 2003.

[34] Krivov, G. G., Shapovalov, M. V., and Dunbrack, R. L. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins*, 77(4):778–795, December 2009.

[35] Mackerell, A. D., Feig, M., and Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry*, 25(11):1400–1415, August 2004.

[36] MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry. B*, 102(18):3586–3616, April 1998.

[37] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, July 1983.

[38] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., KalÃľ, L., and Schulten, K. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802, December 2005.

[39] Brooks, B., Brooks, C., MacKerell, A., Nilsson, L., Petrella, R., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R., Post, C., Pu, J., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D., and Karplus, M. CHARMM: The Biomolecular Simulation Program. *Journal of computational chemistry*, 30(10):1545–1614, July 2009.

[40] Krautler, V., van Gunsteren, W. F., and Hunenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *Journal of Computational Chemistry*, 22(5):501–508, April 2001.

[41] Tuckerman, M., Berne, B. J., and Martyna, G. J. Reversible multiple time scale molecular dynamics. *The Journal of Chemical Physics*, 97(3):1990–2001, August 1992.

[42] Shan, Y., Klepeis, J. L., Eastwood, M. P., Dror, R. O., and Shaw, D. E. Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *The Journal of Chemical Physics*, 122(5):54101, February 2005.

[43] Shaw, D. E., Grossman, J. P., Bank, J. A., Batson, B., Butts, J. A., Chao, J. C., Deneroff, M. M., Dror, R. O., Even, A., Fenton, C. H., Forte, A., Gagliardo, J., Gill, G., Greskamp, B., Ho, C. R., Ierardi, D. J., Iserovich, L., Kuskin, J. S., Larson, R. H., Layman, T., Lee, L. S., Lerer, A. K., Li, C., Killebrew, D., Mackenzie, K. M., Mok, S. Y. H., Moraes, M. A., Mueller, R., Nociolo, L. J., Peticolas, J. L., Quan, T., Ramot, D., Salmon, J. K., Scarpazza, D. P., Schafer, U. B., Siddique, N., Snyder, C. W., Spengler, J., Tang, P. T. P., Theobald, M., Toma, H., Towles, B., Vitale, B., Wang, S. C., and Young, C. Anton 2: Raising the Bar for Performance and Programmability in a Special-Purpose Molecular Dynamics Supercomputer. In *SC14: International Conference for High Perfor-*

mance Computing, Networking, Storage and Analysis, pages 41–53, November 2014.

[44] Zhu, F. and Hummer, G. Convergence and error estimation in free energy calculations using the weighted histogram analysis method. *Journal of Computational Chemistry*, 33(4):453–465, February 2012.

[45] Alberty, R. A. and Hammes, G. G. Application of the Theory of Diffusion-controlled Reactions to Enzyme Kinetics. *The Journal of Physical Chemistry*, 62(2):154–159, February 1958.

[46] Chou, K. C. and Zhou, G. P. Role of the protein outside active site on the diffusion-controlled reaction of enzymes. *Journal of the American Chemical Society*, 104(5):1409–1413, March 1982.

[47] Wade, R. C., Gabdoulline, R. R., Ludemann, S. K., and Lounnas, V. Electrostatic steering and ionic tethering in enzyme-ligand binding: Insights from simulations. *Proceedings of the National Academy of Sciences*, 95(11):5942–5949, May 1998.

[48] Robert, A. and Howe, J. R. How AMPA receptor desensitization depends on receptor occupancy. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 23(3):847–858, February 2003.

[49] Durr, K. L., Chen, L., Stein, R. A., De Zorzi, R., Folea, I. M., Walz, T., Mchaourab, H. S., and Gouaux, E. Structure and dynamics of AMPA receptor GluA2 in resting, pre-open, and desensitized states. *Cell*, 158(4):778–792, August 2014.

Chapter 5 - Energetics of glutamate binding to an AMPA receptor

Reproduced in part from a manuscript submitted to:

The Journal of Physical Chemistry B

by

Alvin Yu and Albert Y. Lau

Ionotropic glutamate receptors (iGluRs) are ligand-gated ion channels that are responsible for the majority of excitatory transmission at the synaptic cleft. Mechanically speaking, agonist binding to the ligand binding domain (LBD) activates the receptor by triggering a conformational change that is transmitted to the transmembrane region, opening the ion channel pore. We use fully atomistic molecular dynamics simulations to investigate the binding process in the AMPA receptor, an iGluR subtype. The string method with swarms of trajectories was applied to calculate the possible pathways glutamate traverses during ligand binding. Residues peripheral to the binding cleft are found to metastably bind the ligand prior to ligand entry into the binding pocket. Umbrella sampling simulations were performed to compute the free energy barriers along the binding pathways. The calculated free energy profiles demonstrate that metastable interactions contribute substantially to the energetics of ligand binding and form local minima in the overall free energy landscape. Protein ligand interactions at sites outside of the orthosteric agonist-binding site may serve to lower the transition barriers of the binding process.

5.1 Background

In the brain, glutamate binding serves as an essential signal for activation of downstream neurons [1]. AMPA receptors, composed of an amino-terminal domain (ATD), a ligand binding domain (LBD), a transmembrane domain (TMD) and C-terminal domain (CTD), sense the release of neurotransmitters at synaptic terminals [2]. Architecturally, AMPA receptors are tetrameric complexes arranged in a dimer-of-dimers fashion [3]. The binding of neurotransmitters triggers large structural rearrangements in which conformational change in the separate LBDs provides the necessary tension for opening the channel pore. Isolated LBDs close once glutamate binds. Thus, ligand binding and LBD closure are thought to be tightly coupled processes.

Many computational methods have been employed to study protein ligand binding. Unbiased simulation [4, 5] is the most straightforward approach and yields detailed information on the molecular mechanisms of the binding process, but it is also the most computationally expensive, often requiring the use of special purpose hardware

[6]. Usually, because of the computational cost involved, only a limited number of binding events are sampled. Enhanced sampling methods, such as umbrella sampling and the string method, are an alternative approach.

For the AMPA receptor subtype, GluA2, umbrella sampling has been used to calculate the free energy of ligand binding for multiple agonists [7]. The energetics of cleft closure and ligand docking were evaluated separately. The contributions to the free energy from each was computed along a predetermined pathway, using biasing potentials to pull the system from a ligand free, open LBD conformation to a ligand bound, closed LBD conformation.

Here, we aim to re-examine glutamate receptor ligand binding using the string method [8, 9]. In contrast to our previous study [7], translation of the ligand into the binding pocket is not restricted to a predetermined path, and ligand translation and protein conformational change are studied as coupled processes in concert. As a result, we find ligand-binding pathways that are physically more realistic and putative protein-ligand interactions that occur outside of the binding pocket.

5.2 Computational Modeling and Methods

5.2.1 Simulation System Preparation.

The atomic model for the ligand bound monomer was constructed from the crystal structure of the GluA2 ligand-binding core (S1S2) in complex with glutamate (PDB ID: 1FTJ). The atomic model for the ligand free monomer was constructed from the crystal structure of the GluA2 ligand-binding core (S1S2) in the apo state (PDB ID: 1FTO). Missing amino acid residues backbones were added to the model using the Modloop server [10], and missing sidechains were constructed using SCWRL4 [11]. Models included the crystallographic waters from the binding cleft, and both systems were solvated with 13,836 water molecules. A glutamate molecule was added to bulk solvent in the apo system ∼38 Å away from the LBD center of mass. Both systems were neutralized by adding Na^+ and Cl^- ions to the bulk solution until the salt concentration reached 150 mM. Both systems contained a total of 45,723 atoms. Pe-

riodic boundary conditions were imposed on an orthorhombic cell with approximate dimensions of 84 Å × 68 Å × 78 Å. The systems were energy minimized and equilibrated at constant pressure and temperature (NPT) conditions at 1 atm and 300K with a time step of 2 fs. For all simulations, the all-atom potential energy function PARAM27 [12,13] for proteins and the TIP3P potential energy function for water [14] were used. Electrostatic interactions were computing using the particle mesh Ewald (PME) algorithm and short-range, non-bonded interactions were truncated at 12 Å. All productions runs were performed in CHARMM [15].

5.2.2 String Method Calculation.

Conformational transitions in biological macromolecules are multi-dimensional processes that require the concerted movement along a large number of degrees of freedom. The important order parameters that well describe these transitions are often not known a priori. The string method is a "chain-of-states" approach that optimizes the transition states between two end points and has been used to characterize conformational change in tyrosine kinases, ion channels, and motor proteins [16–19].

Briefly, the string method interpolates between two end states by sampling states intermediate to the transition [8, 9]. The algorithm consists of first constructing an initial interpolation between the two end states to form a transition pathway. The pathway is then iteratively refined as follows until it converges: (1) each transition state is evolved using molecular dynamics simulations to estimate the instantaneous forces underlying the transition; (2) states are moved in the direction of those forces; (3) the entire pathway is reparametrized to maintain some metric to separate the states - typically this metric is taken to be the Euclidean distance in the space of collective variables. The two end states used in our study were taken from the prepared atomic models of the ligand-bound and ligand-free LBD monomers.

Three initial transition pathways were constructed based on the accessibility of the ligand to the binding pocket. In the interpolation for pathway 1, the ligand was translated into the binding cleft between K449 and S652 using biasing potentials. For the interpolations in pathways 2 and 3, glutamate was translated into the binding

pocket via the ξ_1 and ξ_2 sides of the cleft respectively, where ξ_1 and ξ_2 are two center of mass distances between lobe 1 and lobe 2 (**Fig. 5-1**), using biasing potentials. The total number of states or images for pathways 1, 2, and 3 were 129, 132 and 129, respectively. For the string method calculations our collective variable was the Cartesian coordinates of all Cα atoms of the LBD, and the two terminal carbonyl carbons and the Cα atom of the ligand.

The string method algorithm with swarms of trajectories [9] was applied to all three pathways. Pathways 1 and 2 converged after ∼225 ns of aggregate sampling. Pathway 3, however, did not converge after 258 ns of sampling and was discarded for possibly being pathological. Timesteps of 2.0 fs and 1.0 fs were used in the free and restrained simulations, respectively. Each iteration of the string method involved the following: (1) 5,000 steps of restrained dynamics to push the pathway to the target collective variable values; (2) a short restrained dynamics run for 500 steps to generate the starting coordinates for the swarms; starting coordinates were written out at 5 step intervals; (3) 100 simulations of short 50-step unbiased simulations to measure the average drift of the swarms; (4) the average drift was measured, and the positions of the images were updated using the average drift; (5) the path was reparameterized, and the target collective variables were updated for the next iteration with the current position of the images. Convergence of the transition paths were assessed by summing the distances between the images at the j-th iteration and the images at the $(j+1)$ iteration, i.e., $c_j = \sum_{i=1}^{N-1} \sqrt{(\vec{\theta}_{i,j} - \vec{\theta}_{i,j+1})^2}$, where $\vec{\theta}_{i,j}$ is a multi-dimensional vector representing the collective variables of image i at the j-th iteration.

5.2.3 The Free Energies of the Transition Paths.

Once a pathway is found, part of the difficulty in capturing the free energy along a particular binding pathway stems from the high dimensionality of the ligand binding process. Protein-ligand interactions, ligand orientation, protein conformational change, and specific residue-residue contacts may each contribute to the free energy along different degrees of freedom [20]. To address this issue, we use umbrella sampling and choose a global order parameter, α, that measures the progress along the

binding process. In the discrete case, as in our present study, α is an index over the states of the system 6defined by the transition states of a converged pathway identified by the string method. Ligand association can then be represented as a transition: $\alpha_N \rightarrow \alpha_{N-1} \rightarrow \cdots \rightarrow \alpha_1 \rightarrow \alpha_0$; where α_N corresponds to an initial state containing the protein and ligand separated by bulk solvent, and α_0 corresponds to a final state containing the ligand bound complex.

The parameter α may depend on many degrees of freedom that capture the conformation of the protein and ligand at the particular transition state of interest. Given the atomic coordinates of a protein-ligand system, how do we determine α and how far along the binding process a system is? In the present study, we compare the Cartesian coordinates of the Cα atoms in the protein to capture the protein conformation and residue specific interactions, and two terminal carbonyl carbons as well as the Cα atom in the ligand to specify ligand orientation and protein-ligand interactions. The system is then assigned to the most similar state, α_i, along the 1-dimensional transition pathway. This is done automatically by first measuring the coordinates of interest at each α_i, Voronoi tessellating the phase space of those coordinates, and assigning each conformation in the neighborhood of α_i to the index i.

The free energy along the binding pathway can then be calculated by umbrella sampling. Voronoi tessellation has been previously used to separate the transition states [21–23]. Here, the same approach is taken, and statistics on the occupancies of each state are collected on a confined region of the 1-dimensional pathway through the use of biasing potentials. The biasing potentials restrain the system along the multi-dimensional coordinates of interest. Contributions of the biasing potential to the free energy can be unbiased using the Weighted Histogram Analysis Method (WHAM) [24, 25]. Thus, the method calculates a free energy profile across a multi-dimensional surface. The accuracy of this PMF depends on the degree of sampling between each state, α_i, and its neighbors.

Umbrella sampling simulations, totaling 105 ns along pathway 1 and 103 ns along pathway 2, were performed to calculate the free energy profiles. Each image was harmonically restrained with force constants (0.01 kcal/mol/Å^2) at each of the collective

variables from the string method. Biased distributions were collected across all collective variables. The phase space of the collective variables was voronoi tessellated, using each image as the center for the voronoi cell. The biased distribution of states along the binding pathway was calculated by assigning systems in the neighborhood of state i to state i based on the voronoi tessellation, i.e.,

$$x \in \alpha_i \leftrightarrow \|\vec{\theta}(x) - \vec{\theta}(\alpha_j)\| \geq \|\vec{\theta}(x) - \vec{\theta}(\alpha_i)\|, \ \forall j \neq i$$

where x represents a conformation of the system collected during the umbrella sampling, α_i is the system at image index i, and $\vec{\theta}(x)$ is a multidimensional vector containing the collective variable values of conformation x. The biased distributions were unbiased using WHAM.

5.3 Results and Discussion

5.3.1 Local optimization of metastable binding intermediates.

Binding along converged pathways 1 and 2 contains distinct interactions with residues on the periphery of the LBD cleft that are not directly in the binding pocket. The presence of multiple binding pathways suggests that ligand binding is not restricted to a single pathway. In pathway 1, the ligand forms interactions with R453 and R661, whereas in pathway 2, the ligand forms interactions with R684 and K449. In particular, R453 and R684 pass the ligand to R660 and K449, respectively, prior to binding. R660 and K449 form metastable interactions with the ligand before passing it into the pocket (**Fig. 5-2** C-E; **Fig. 5-3** C-E). Once inside the pocket, the ligand contacts Y450 and R485 on lobe 1, and E705 on lobe 2, as the LBD closes.

Certain residues were observed to be involved in locking the LBD closed. A salt bridge between K730 and D728 in the hinge region of the LBD forms after the ligand settles into the binding pocket. The interaction between K730 and D728 may be correlated with the degree of cleft closure. This salt bridge forms once the LBD is closed to (ξ_1, ξ_2) = (11.9, 11.4 Å) in pathway 1 and to (ξ_1, ξ_2) = (9.5, 5.3 Å) in pathway 2. Previous crystallographic experiments showed K730 forms a salt bridge with E705

in the apo state; however, when a ligand is bound, K730 switches interaction partners to D728 [26]. Also consistent with the crystal structures, the D651-S652 peptide bond flips to form additional hydrogen bonding interactions between lobe 1 and lobe 2 of the LBD; the backbone carbonyl group of S652 (lobe 2) hydrogen bonds with the backbone amide of G451 (lobe 1).

5.3.2 Characterizing the energetics of molecular interactions along a binding pathway.

We computed the free energy profile along each of the converged binding pathways obtained with the string method. In pathway 1 (**Fig. 5-4**), the free energy plateaus when the protein and ligand are separated by ∼12 Å of solvent ($i = 90$) and oscillates around 7-8 kcal/mol as the electrostatic effects attracting the negatively charged ligand and positive residues in the binding cleft decay. Contact between the ligand and specific residues on the periphery of the binding cleft lower this free energy barrier. R684 contacts the ligand's α-carboxylate ($i = 73$), lowering the free energy by ∼2 kcal/mol, and the S652 side chain hydrogen bonds with the ligand ($i = 61$), lowering the free energy by ∼1 kcal/mol. Notably, K449 contacts the ligand ($i = 50$) and passes it into the binding pocket. As the ligand enters from the ξ_1 side of the binding cleft, it interacts with three separate residues that stabilize a metastable transition state in which (1) the ligand's α- carboxylate coordinates the amide group of Y450, (2) the ligand's amide group interacts with the hydroxyl side chain of S652, and (3) the ligand's γ-carboxylate contacts K449 (i =32). This metastable interaction decreases the free energy barrier by ∼3 kcal/mol. The α-carboxylate shifts into the binding pocket to contact R485 (i =27), whereas the γ-carboxylate rotates downward to contact S654 and T655 ($i = 13$). Interactions that lock the LBD closed, including lobe 1 - lobe 2 interactions between S652 and G451 ($i = 24$) and the salt bridge between K730 and D728 in the hinge region ($i = 14$), raise the free energy by ∼2 kcal/mol.

Similarly in pathway 2 (**Fig. 5-5**), attractive protein ligand interactions plateau at ∼10 Å ($i = 102$). The ligand first contacts the LBD at R453 on the ξ_1 side, lowering

the free energy barrier by ∼5 kcal/mol. R661 contacts the γ-carboxylate ($i = 29$), decreasing the free energy by ∼6 kcal/mol. R485 rotates out of the binding pocket via rotations of +130, +20, -124, and -61 around the R485 sidechain's χ_1, χ_2, χ_3, and χ_4 torsion angles, respectively (**Fig. 5-5** F, K) to coordinate the γ-carboxylate ($i = 17$). Glutamate moves into the binding pocket while its γ- carboxylate is coordinated by R485, and then flips downward so that the γ-carboxylate contacts S654, whereas the α-carboxylate flips up to contact R485 ($i = 13$). The LBD locks closed as the D651-S652 peptide bond flips ($i = 13$) and the K730-D728 salt bridge forms ($i = 8$). These interactions that close the LBD do not raise the free energy as in pathway 1; however, the final ligand-bound complex ($i = 0$) increases the free energy by ∼3 kcal/mol.

The total binding free energy is given by: $\Delta G_{\text{bind}} = \Delta G_0 + \Delta G_N + \sum_{i=0}^{N-1} \Delta G_{i,i+1}$, where $\Delta G_{i,i+1}$ is the relative free energy difference between the system at state i and state $i + 1$. ΔG_0 is the free energy difference between the initial state and the unbiased ensemble of bound protein- ligand conformations. ΔG_N, likewise, is the free energy difference between the final state and the unbiased ensemble of apo protein and free ligand conformations. Assuming the initial ($i = 0$) and final ($i = \text{N}$) states of the binding pathway are well sampled, ΔG_0 and ΔG_N may be small. Alternatively, ΔG_0 and ΔG_N may be calculated by methods that estimate the free energy of restraining the orientational, conformational, and translational change in both the protein and ligand during binding [7].

The overall free energy difference between the initial and final state, $\sum_{i=0}^{N-1} \Delta G_{i,i+1}$, is -5.8 kcal/mol in pathway 1 and -8.8 kcal/mol in pathway 2, close to the experimentally measured binding free energy of -8.3 kcal/mol (IC 50) derived from competition assays of radiolabeled AMPA with glutamate [26]. Discrepancies in the calculated free energy values may be due to differences in the initial and final protein-ligand conformations between the two pathways, undersampling of state-to-state transitions, and inherent difficulties of free energy estimation in high dimensional spaces. The free energy difference between the end points may be compared to the experimental binding free energies if these discrepancies are small and the end points are representative of

the ligand bound and ligand free ensembles.

5.4 Conclusions

In the present study, we have utilized a "chain-of-states" approach to probe glutamate binding in the AMPA receptor. The binding pathways converged after a few hundred nanoseconds of simulation time, whereas unbiased simulations typically require simulation times on the order of tens of microseconds to sample binding events in proteins which have similar on rates ($k_{\text{on}} \sim 10^7$ $\text{M}^{-1}\text{s}^{-1}$) [5]. Metastable interactions between the ligand and positively charged residues on helix F (R660, R661), R684, K660, and R453 lower the free energy barrier during ligand binding, and form local minima in the free energy landscape. The converged glutamate binding pathways indicate that the ligand binds either via the ξ_1 side of the LBD or the space formed between K449 and S652.

The string method approach has several limitations. Most notably, because the algorithm only moves images down an energy gradient, intermediate states can fall into local minima and become "trapped", requiring several different initial interpolations to sample multiple transition pathways. It is difficult to sample all possible transition pathways using the string method, and other low energy pathways for glutamate binding may exist, distinct from the two observed pathways.

Free energy profiles calculated along the binding pathways demonstrate that specific residue- ligand interactions outside of the binding pocket contribute substantially to the energetics of ligand binding. Why might a protein domain contain such low affinity binding sites that do not engage the ligand in its stably bound form? We speculate that these interactions may serve to position the ligand into conformations that are more competent to bind. If this is true, metastable binding sites can facilitate the diffusion of ligands across narrow, tight spaces, like the recessed binding pocket of the AMPA receptor LBD. In general, proteins have adopted a wide range of strategies to transfer molecules to locations where they can be acted upon, so it is not implausible that low affinity binding sites situated in strategically positioned locations may be another general feature for molecular transport.

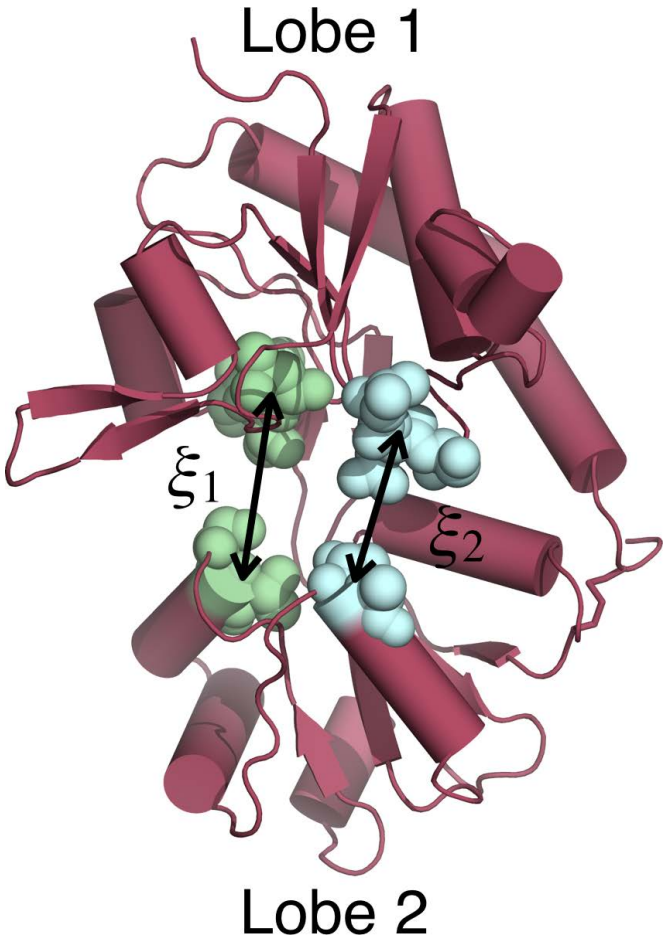

Fig. 5-1: **Opening and closing of the GluA2 LBD is described by the two-dimensional order parameter (ξ_1, ξ_2).** ξ_1 and ξ_2 each indicate the center-of-mass distance between the atoms shown in green and cyan, respectively.

Fig. 5-2 Locally optimized binding intermediates in pathway 1. (A) The initial conformation used in the string method for pathway 1 contains an open LBD (ξ_1, ξ_2) = (14.1 Å, 13.9 Å) and a ligand separated by bulk solvent. (B) Close-up view of (A). (C) Glutamate contacts R684 on lobe 2. (D) The S652 hydroxyl sidechain hydrogen bonds with the amide nitrogen of glutamate. (E) K449 interacts with the ligand's γ-carboxyl (F) Glutamate shifts into the binding pocket. (G) Glutamate forms thee contacts across lobe 1 and lobe 2 of the LBD: (1) The α-carboxyl contacts the amide backbone of Y450. (2) The amide nitrogen interacts with the S652 hydroxyl side chain. (3) K449 contacts the γ-carboxyl. (H) Glutamate is coordinated by the aromatic sidechain of Y450, aligning with the α-carboxyl proximal to R485. (I) The α-carboxyl contacts R485 on lobe 1 and E705 on lobe 2. (J-K) The LBD closes further around the ligand (ξ_1, ξ_2) = (9.5 Å, 8.3 Å) as the γ-carboxyl contacts the amide backbones of S654 and T655. (L) Expanded view of (K).

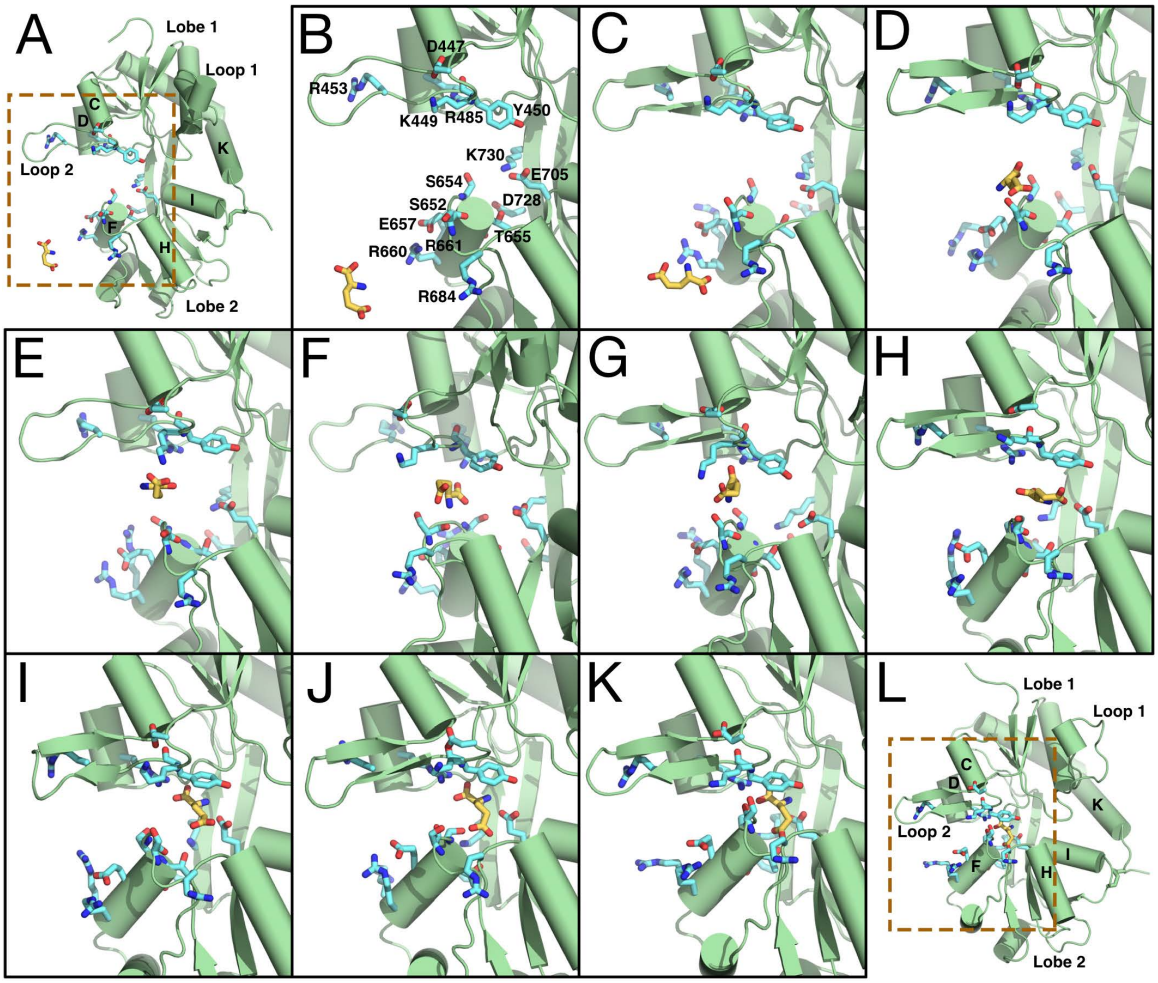
A
Lobe 1
Loop 1
C
D
K
Loop 2
I
F
H
Lobe 2
B
D447
R453
K449
R485
Y450
K730
E705
S654
D728
S652
E657
T655
R661
R660
R684
C
D
E
F
G
H
I
J
K
L
Lobe 1
Loop 1
C
D
K
Loop 2
F
H
I
Lobe 2

Fig. 5-2

Fig. 5-3 Locally optimized binding intermediates in pathway 2. (A) The initial conformation used in the string method for pathway 2 contains an open LBD (ξ_1, ξ_2) = (14.1 Å, 13.9 Å) and a ligand separated by bulk solvent. (B) Close-up view of (A). (C-D) Glutamate contacts R453 on the ξ_1 side of lobe 1. (E) R661 contacts the ligand's γ-carboxyl (F) R485 flickers out of the binding pocket to contact the γ-carboxyl. (G-I) R485 relaxes toward the binding pocket while coordinating the ligand's γ-carboxyl. (J) The γ-carboxyl flips downward into the binding pocket, while the α-carboxyl flips upward to contact R485. The amide nitrogen contacts E705 on lobe 2. (K) The LBD closes further around the ligand (ξ_1, ξ_2) = (9.5 Å, 8.3 Å) as the γ-carboxyl coordinates the amide backbones of S654 and T655. (L) Expanded view of (K).

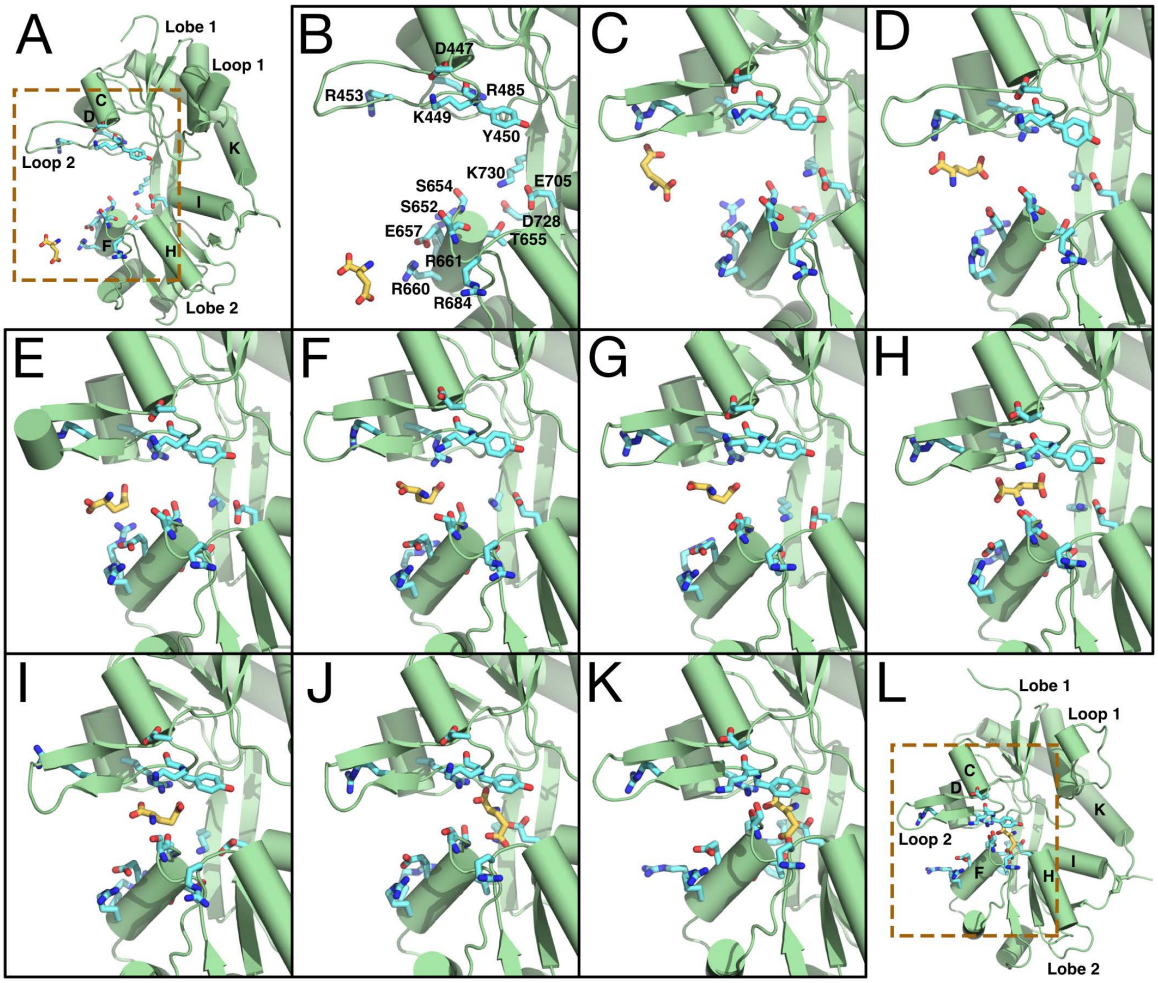
A
Lobe 1
Loop 1
C
D
K
Loop 2
I
F
H
Lobe 2
B
D447
R453
R485
K449
Y450
K730
E705
S654
S652
D728
E657
T655
R661
R660
R684
C
D
E
F
G
H
I
J
K
L
Lobe 1
Loop 1
C
D
K
Loop 2
F
H
I
Lobe 2

Fig. 5-3

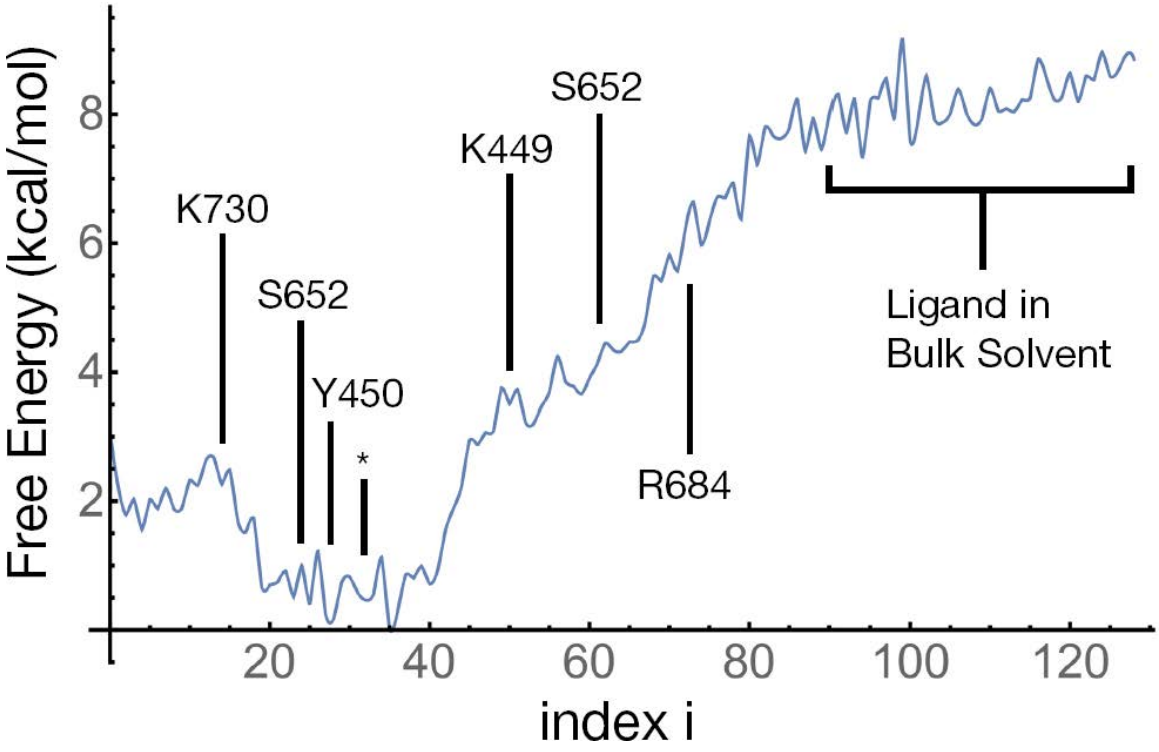

Fig. 5-4: **The free energy profile along binding pathway 1.** The indices refer to images or conformations along pathway 1, calculated by the string method. The free energy was calculated by voronoi tessellating the phase space of the collective variables and performing umbrella sampling. At images (i = 90 - 129) the ligand is in bulk solvent, and at image 0, the ligand is fully bound within the LBD, in the crystallographic conformation. Protein-ligand interactions that change the energetics of ligand binding are labeled. The asterisk (*) indicates an image that contains a metastable intermediate consisting of the following interactions: ligand α-carboxyl to Y450 amide backbone, ligand amide to S652 hydroxyl side chain, and ligand γ-carboxyl to K449 sidechain.

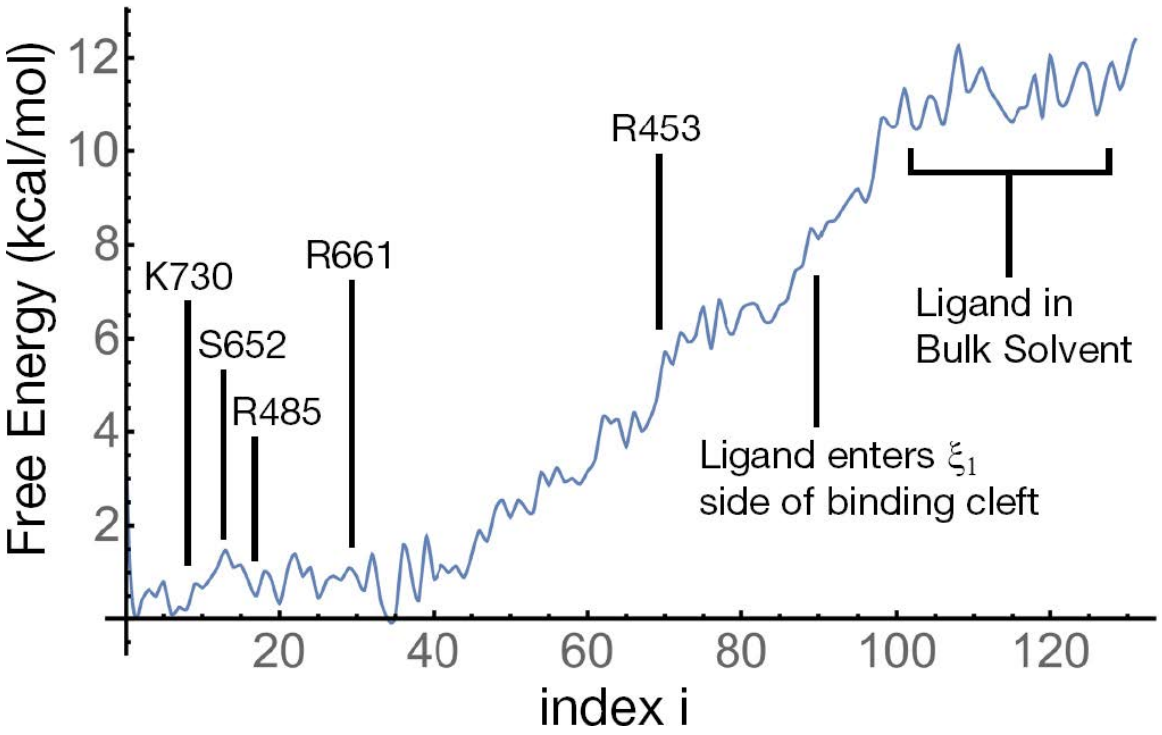

Fig. 5-5: **The free energy profile along binding pathway 2.** The indices refer to images or conformations along pathway 2, calculated by the string method. The free energy was calculated by voronoi tessellating the phase space of the collective variables and performing umbrella sampling. At images (i = 102âĂŞ131) the ligand is in bulk solvent, and at image 0, the ligand is fully bound within the LBD, in the crystallographic conformation. Protein-ligand interactions that change the energetics of ligand binding are labeled.

References

[1] Traynelis, S. F., Wollmuth, L. P., McBain, C. J., Menniti, F. S., Vance, K. M., Ogden, K. K., Hansen, K. B., Yuan, H., Myers, S. J., and Dingledine, R. Glutamate receptor ion channels: structure, regulation, and function. *Pharmacological Reviews*, 62(3):405–496, September 2010.

[2] Mayer, M. L. Glutamate receptors at atomic resolution. *Nature*, 440(7083):456–462, March 2006.

[3] Mayer, M. L. Structure and mechanism of glutamate receptor ion channel assembly, activation and modulation. *Current opinion in neurobiology*, 21(2):283–290, April 2011.

[4] Shan, Y., Kim, E. T., Eastwood, M. P., Dror, R. O., Seeliger, M. A., and Shaw, D. E. How Does a Drug Molecule Find Its Target Binding Site? *Journal of the American Chemical Society*, 133(24):9181–9183, June 2011.

[5] Dror, R. O., Pan, A. C., Arlow, D. H., Borhani, D. W., Maragakis, P., Shan, Y., Xu, H., and Shaw, D. E. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proceedings of the National Academy of Sciences*, 108(32):13118–13123, August 2011.

[6] Shaw, D. E., Dror, R. O., Salmon, J. K., Grossman, J. P., Mackenzie, K. M., Bank, J. A., Young, C., Deneroff, M. M., Batson, B., Bowers, K. J., Chow, E., Eastwood, M. P., Ierardi, D. J., Klepeis, J. L., Kuskin, J. S., Larson, R. H., Lindorff-Larsen, K., Maragakis, P., Moraes, M. A., Piana, S., Shan, Y., and Towles, B. Millisecond-scale Molecular Dynamics Simulations on Anton. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*, SC '09, pages 39:1–39:11, New York, NY, USA, 2009. ACM.

[7] Lau, A. Y. and Roux, B. The hidden energetics of ligand binding and activation in a glutamate receptor. *Nature Structural & Molecular Biology*, 18(3):283–287, March 2011.

[8] Maragliano, L., Fischer, A., Vanden-Eijnden, E., and Ciccotti, G. String method in collective variables: Minimum free energy paths and isocommittor surfaces. *The Journal of Chemical Physics*, 125(2):024106, July 2006.

[9] Pan, A. C., Sezer, D., and Roux, B. Finding Transition Pathways Using the String Method with Swarms of Trajectories. *The Journal of Physical Chemistry B*, 112(11):3432–3440, March 2008.

[10] Fiser, A. and Sali, A. ModLoop: automated modeling of loops in protein structures. *Bioinformatics (Oxford, England)*, 19(18):2500–2501, December 2003.

[11] Krivov, G. G., Shapovalov, M. V., and Dunbrack, R. L. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins*, 77(4):778–795, December 2009.

[12] Mackerell, A. D., Feig, M., and Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry*, 25(11):1400–1415, August 2004.

[13] MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *The Journal of Physical Chemistry B*, 102(18):3586–3616, April 1998.

[14] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, July 1983.

[15] Brooks, B., Brooks, C., MacKerell, A., Nilsson, L., Petrella, R., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R., Post, C., Pu, J., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D., and Karplus, M. CHARMM: The Biomolecular Simulation Program. *Journal of computational chemistry*, 30(10):1545–1614, July 2009.

[16] Meng, Y., Shukla, D., Pande, V. S., and Roux, B. Transition path theory analysis of c-Src kinase activation. *Proceedings of the National Academy of Sciences*, 113(33):9193–9198, August 2016.

[17] Zhu, F. and Hummer, G. Pore opening and closing of a pentameric ligand-gated ion channel. *Proceedings of the National Academy of Sciences of the United States of America*, 107(46):19814–19819, November 2010.

[18] Lev, B., Murail, S., Poitevin, F., Cromer, B. A., Baaden, M., Delarue, M., and Allen, T. W. String method solution of the gating pathways for a pentameric ligand-gated ion channel. *Proceedings of the National Academy of Sciences of the United States of America*, 114(21):E4158–E4167, May 2017.

[19] Ovchinnikov, V., Cecchini, M., Vanden-Eijnden, E., and Karplus, M. A Conformational Transition in the Myosin VI Converter Contributes to the Variable Step Size. *Biophysical Journal*, 101(10):2436–2444, November 2011.

[20] Chodera, J. D., Mobley, D. L., Shirts, M. R., Dixon, R. W., Branson, K., and Pande, V. S. Alchemical free energy methods for drug discovery: Progress and challenges. *Current opinion in structural biology*, 21(2):150–160, April 2011.

[21] Dickson, A., Warmflash, A., and Dinner, A. R. Nonequilibrium umbrella sampling in spaces of many order parameters. *The Journal of Chemical Physics*, 130(7):074104, February 2009.

[22] Vanden-Eijnden, E. and Venturoli, M. Markovian milestoning with Voronoi tessellations. *The Journal of Chemical Physics*, 130(19):194101, May 2009.

[23] Maragliano, L., Vanden-Eijnden, E., and Roux, B. Free energy and kinetics of conformational transitions from Voronoi tessellated milestoning with restraining potentials. *Journal of Chemical Theory and Computation*, 5(10):2589–2594, August 2009.

[24] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, 13(8):1011–1021, October 1992.

[25] Souaille, M. and Roux, B. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer Physics Communications*, 135:40–57, March 2001.

[26] Armstrong, N. and Gouaux, E. Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron*, 28(1):165–181, October 2000.

Chapter 6 - Concluding Remarks

There are many unanswered questions regarding ionotropic glutamate receptor behavior. In this dissertation, we have focused on understanding two parts of iGluR activation - ligand binding, and LBD conformational change. In **Chapter 2**, we outlined the umbrella sampling free energy method we used to understand the energetics of these processes. **Chapter 3** examined the free energy landscapes of conformational change in a particularly high affinity glycine binding domain. **Chapter 4** identified metastable binding sites which funnel the ligand into the binding pocket and presented molecular simulations of ligand binding in iGluRs. **Chapter 5** uncovered multiple possible ligand binding pathways and examined the energetics of those pathways.

Notably lacking in these studies, is a rigorous understanding of the pore opening, ion conduction, and desensitization mechanisms of these receptors. Future studies regarding these events will likely require molecular simulations of the complete, intact receptor and would enhance our understanding of the physical and chemical basis for neural activity.

ALVIN YU

PERSONAL INFORMATION

address	23 Pierside Dr. Unit 320 Baltimore, MD 21230
email	alvin.yu@jhu.edu
phone	(M) +1 (626) 283 2924

PERSONAL STATEMENT

My current research is on atomistic-level descriptions of ionotropic glutamate receptor behavior, performing molecular dynamics simulations on both iGluR domains (∼40,000 atoms) and large intact complexes of the receptor (∼1 million atoms). A variety of computational methods are used to probe iGluR ligand binding, conformational change, and activation. These methods include, but are not limited to: long-timescale simulation, free energy calculations and enhanced sampling.

EDUCATION

2011–Present Johns Hopkins University

PhD in Biophysics

Department of Biophysics & Biophysical Chemistry
Thesis: *Neurotrasmitter funneling optimizes glutamate receptor kinetics*
Description: Characterized a primary mechanism of how neurons sense chemical signals at the synapse during activation of neural impulses, through a combination of computationally intensive simulations on supercomputing resources (e.g. ANTON) and design of electrophysiology experiments (in collaboration with AJR Plested). These findings can help guide the discovery of new therapeutic targets for glutamate receptor dysfunction (e.g. Alzheimer's).
Advisor: Prof. Albert. Y. Lau

2006–2011 California Institute of Technology

Bachelors in Physics

GPA: 4.0 · Division of Physics Mathematics & Astronomy
Description: This degree focused on general principles in physics, mathematics, and science, providing a rigorous foundation for quantitative research.

RESEARCH EXPERIENCE

2011–Present Graduate Researcher – School of Medicine

Johns Hopkins University

Spearheaded a collaboration with the Mayer lab at the NIH and the Plested lab at the Leibniz-Institute für Molekulare Pharmakologie, including both design of experiments and providing computational/theoretical expertise. Mentored graduate students and undergraduate students, helping to formulate research plans and projects.

2010–2011 Researcher – Condensed Matter Physics Dept.

California Institute of Technology

Designed and wrote programs to simulate the interactions between hard-core bosons on a two-dimensional lattice model of high-Tc superconductors. Worked on algorithmic improvements, using symmetry in the system to speed up computational bottlenecks on a Beowulf cluster. Analyzed and compared data to experimental observations, verifying previously known scaling relations (i.e. Homes's Law).

Jun-Aug 2008 Summer Intern – Jet Propulsion Laboratory

NASA

Devised a method to analyze the seasonal variation in the mass and density of surficial frost on the Martian Northern Polar Cap. Implemented automated pattern recognition of surface reflections and filtered out noise fluctuations in radar data taken from the Shallow RADar (SHARAD) instrument on board the Mars Reconaissance Orbiter (MRO) satellite.

PUBLICATIONS

Yu A, Salazar H, Plested AJR, and Lau AY. Neurotrasmitter funneling optimizes glutamate receptor kinetics (*Submitted*)

Yu A, and Lau AY. Energetics of glutamate binding to an AMPA receptor (*Submitted*)

Yu A, Albertstein A, Thomas A, Zimmer A, Grey R, Mayer ML and Lau AY. Molecular lock regulates binding of glycine to a primitive NMDA receptor *Proc Natl Acad Sci USA* 113(44):E6786-E6795

Yu A, Wied T, Belcher J, and Lau AY. Computing conformational free energies of iGluR ligand binding domains. In: Popescu GK, editor. Ionotropic glutatmate receptor technologies. New York: Humana Press;

SKILLS

Technical Skills
Molecular Dynamics (e.g. CHARMM, NAMD)
Free Energy Calculations (e.g. umbrella sampling)
Scientific Computing (e.g. MATLAB, Mathematica)
Programming (e.g. python, C++, Fortran)
Graphic Design Tools (e.g. Photoshop)
Standard Computing (e.g. Linux)

Core Competencies
Mentorship and Project Leadership
Collaboration and Group Work
Analysis of Scientific/Technical Literature
Effective Communication to Technical and Non-Technical Audiences
Developing/Executing Research

OTHER INFORMATION

Awards

2013-2016 · Thomas J. Kelly and Mary L. Kelly Young Scholar Fund

2011-2013 · Francis D. Carlson Fellowship

2008 · Caltech Summer Undergraduate Research Fellowship

Presentations & Talks

2017 · Probing a molecular lock in a primitive NMDA receptor. In: *61st Annual Meeting of the Biophysical Society.* Presentation/Talk

2016 · Mechanisms of glutamate capture by a clamshell binding domain. In: *60th Annual Meeting of the Biophysical Society.* Poster Abstract

2015 · Simulations of ligand binding in a glutamate receptor. In: *The 24th Annual Institute of Biophysical Research Retreat.* Presentation/Talk

2015 · Long timescale simulations of ligand binding in glutamate receptors. In: *59th Annual Meeting of the Biophysical Society.* Poster Abstract

2014 · Free energy calculations and markov state models in glutamate receptor ligand binding In: *The 58th Annual Meeting of the Biophysical Society* Poster Abstract

2013 · Using the string method to examine ligand binding pathways and transitions in a glutamate receptor. In: *The 57th Annual Meeting of the Biophysical Society* Poster Abstract

Languages

Fluent · English, Mandarin Chinese

Basic · French, Japanese

Interests

Hiking & Camping · Cooking · Running · Simulations · Beer

REFERENCES

Prof. Albert Y. Lau
Dept. of Biophysics & Biophysical Chemistry
Johns Hopkins University School of Medicine
email:alau@jhmi.edu
phone: +1 (443) 187 4528

Prof. Mario Amzel
Dept. of Biophysics & Biophysical Chemistry
Johns Hopkins University School of Medicine
email:mamzel@jhmi.edu
phone: +1 (410) 955 3955

Prof. Mark L. Mayer
National Institutes of Health NICHD DHHS
email:mayerm@mail.nih.gov
phone: +1 (301) 496 9346

Prof. Andrew Plested
Leibniz-Institut für molekulare pharmakologie (FMP)
email:plested@fmp-berlin.de
phone: +49 30 9406 3071 ext 3015

July 25, 2017

161